Monotonicity Properties of Systems of Ordinary Differential Equations

by

Herbert Eduard Kunze

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 1997

The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

A general framework for determining when a solution component is monotone with respect to changes in an initial component value is developed.

Conditions for monotonicity with respect to an orthant are formulated graph theoretically, and conditions for partial strong monotonicity are given.

Monotonicity with respect to a closed, convex cone, K, is also investigated. For a system of differential equations, $\dot{\tilde{x}} = \tilde{f}(\tilde{x})$, $\tilde{x}(0) = \tilde{x}_0$, $\tilde{x} \in \Omega$, the Kamke-Müller Theorem (1932/1927) is extended to closed, convex cones by imposing the essential hypothesis

$$\exists\, l \text{ such that } D\tilde{f}(\tilde{x}) + lI : K \mapsto K, \forall\, \tilde{x} \in N, N \text{ compact.}$$

Strong monotonicity is achieved by further demanding that

$$\exists\, m \text{ such that } (D\tilde{f}(\tilde{x}) + (l+1)I)^m : K \setminus \{\tilde{0}\} \mapsto int(K), \forall\, \tilde{x} \in N,$$

or, more practically, through a graph theoretic formulation. Given a cone with n generators, $\tilde{e}_i$, a directed multigraph on n vertices, g_i, is constructed with a directed edge from g_i to g_j, $i \neq j$, if $\tilde{e}_j$ is in the smallest face of the cone containing $(D\tilde{f}(\tilde{x}) + (l+1)I)\tilde{e}_i$, $\forall\, \tilde{x} \in N$. The multigraph being strongly connected is a sufficient condition for strong monotonicity.

The results of this thesis are applicable to general autonomous ODEs, but the examples are drawn mostly from chemical kinetics.

Acknowledgements

Thanks to David Siegel for his guidance, encouragement, and, most of all, plenty of mathematical fun.

Thanks to my parents for their support and encouragement.

Thanks to my wife, Karen, for her interest in my work, for her enthusiastic encouragement and support, and for listening to my ideas.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The goal of this work is to establish a general framework for the investigation of monotonicity properties of autonomous systems of ordinary differential equations. Earlier efforts in [18] used very specific arguments for particular problems to establish monotonicity; there was a limited amount of broadly applicable results.

We are primarily interested in how a solution component to a general autonomous system of ODEs changes when a single initial component is changed. If a solution component with a single changed initial component is always greater (less) than the original solution component, then we say that the component is monotone increasing (decreasing) with respect to changes in that initial component value. Alternatively, one can look at the sign of the partial derivative of a solution component with respect to an initial component value. If this derivative does not change sign, then the component is monotone with respect to changes in the corresponding initial component value. We will typically focus on obtaining this type of derivative result.

Monotonicity results are of interest for several reasons, the simplest of which is

that non-linear ODEs theory has broad application. This work adds the spice of graph theory and convex cones to present some surprisingly rich mathematics.

More practically, monotonicity results can allow one to predict the qualitative behaviour of a solution component relative to that same solution component with a changed initial value in some component; this knowledge can lead to an understanding of the stability of solutions under changes in initial values. Furthermore, monotonicity results can also prove useful when deciding if a given mathematical model correctly represents a physical problem of interest. For example, if examination of a proposed mathematical model does not verify certain monotonicity observed in experiments, one could conclude that the proposed model is in error.

The work is presented in two parts. This introduction is followed by a small subsection discussing the mathematical basics of chemical kinetics; it is followed by the two chapters of the main body.

Chapter 2 deals with monotonicity with respect to an orthant. The Kamke-Müller Theorem (1932/1927) is the key foundation upon which many of the results existing in the literature are built. Kamke-Müller-like results allow us to determine when the partial derivatives of interest are non-negative or non-positive. If possible, we will want strict sign, or strong monotonicity, results; the graph theoretic approach proved essential in the proofs of these results. Indeed, a related result in the literature has an error in its proof, a proof with no graph theoretic component. Examples are distributed throughout the chapter, with lengthier examinations of particular problems of interest appearing at the end.

In Chapter 3, the theory of convex cones is merged with the theory of the orthant, producing a more general approach to the problem of monotonicity with respect to initial conditions. After some preliminaries, the Kamke-Müller Theorem

is extended to convex cones; this forms the foundation for a new group of tools for investigating monotonicity. The essential condition of this new theorem is linked to several other conditions that exist in the literature. Once again, a graph theoretic approach proves essential as a practical tool for determining when there is strong monotonicity. Some seemingly strange ideas, such as expanding convex cones, are introduced and then showcased in the examples. After a brief discussion of conditions for finding useful cones, the bulk of the examples is presented. To highlight the additional power of this more complicated approach, results obtained in the examples of Chapter 1 are improved upon.

We close this introduction by making three notational remarks.

Throughout this work, vectors will be denoted by placing a tilde on top of their letter (for example, $\tilde{x}$).

The letter I will have three different meanings. $I(\tilde{x})$ will denote the positive interval of existence of a solution $\tilde{x}$ (explained in the preliminaries of Chapter 1). The identity matrix of the appropriate dimension at the time of use will also be denoted I. In epidemiological examples, $I(t)$ will represent the infective population at time t. It is expected that the reader will be able to discern, based on context, what role the letter I plays in a particular instance.

N will have two different meanings. $\tilde{N}$, often indexed, will be used to denote normal vectors. N will be used to denote a neighbourhood. Once again, context should clarify which meaning is in effect.

1.1 Chemical Kinetics

In the framework of formal mass action kinetics, a given reaction mechanism is translated into a mathematical model by associating a rate expression to each process in the reaction mechanism. We assume that the Law of Mass Action holds in order to determine these rate expressions.

Definition 1 *The* **Law of Mass Action** *postulates that the rate expression of a reaction depends on the product of the concentrations of reacting species raised to the power of their molecularity.*

[8] provides a detailed discussion of the Law of Mass Action. The mass action type kinetic equation was first derived by Wilhelmy in 1850. We should note that this is a macroscopic theory. Our primitive concept is an elementary reaction; we are not concerned with electrons, atoms, or molecules and their arrangement. We assume that all rates of reactions are positive constants; they could, in reality, be temperature dependent, for example. We also assume that any mixtures are homogeneous; that is, the vat within which the reaction takes place is evenly stirred.

A general chemical reaction mechanism is of the form

$$\sum_{i=1}^{n} \alpha_{il} A_i \xrightarrow{k_l} \sum_{i=1}^{n} \beta_{il} A_i \,, \tag{1.1}$$

where $l = 1, \ldots, m$ labels the reactions, the stoichiometric coefficients α_{il} and β_{il} are non-negative integers, and it is assumed that $\beta_{il} \neq \alpha_{il}$ for some i. Let $x_i(t)$ be the concentration of species A_i at time t. Assuming mass action chemical kinetics gives

$$\dot{x}_i(t) = f_i(\tilde{x}) = \sum_{l=1}^{m} (\beta_{il} - \alpha_{il}) r_l(t) = \sum_{l=1}^{m} k_l (\beta_{il} - \alpha_{il}) \prod_{p=1}^{n} (x_p(t))^{\alpha_{pl}}, \tag{1.2}$$

for $i = 1, \ldots, n$. $r_l(t)$ is the rate expression for reaction l.

There are some technical details that merit mentioning. We use atom-free stoichiometry; for some discussion, see [8], page 26. Concentrations are either positive for all positive time or they are identically zero for all time. In the second case, one can consider a new reaction mechanism which fits the first case. This is detailed in [39], chapter 12. We state this as a fundamental assumption, where $\mathcal{O}$ denotes the open positive orthant and $\bar{\mathcal{O}}$ denotes its closure, the non-negative orthant:

Positivity Assumption: For $\tilde{x}_0 \in \mathcal{S}_0 \subset \bar{\mathcal{O}}$, $\varphi_t(\tilde{x}_0)$, the solution to the chemical kinetics system (1.2) with initial condition $\tilde{x}_0$, satisfies $\varphi_t(\tilde{x}_0) \in \mathcal{O}$ for $t > 0$.

Chapter 2

Monotonicity With Respect to an Orthant

2.1 Preliminaries

The positive orthant in $\mathbf{R}^n$, denoted $\mathcal{O}$, is given by $\{x_i | x_i > 0, i = 1, \ldots, n\}$. The non-negative orthant is denoted $\bar{\mathcal{O}}$. The results in this chapter will often involve inequalities between vectors. Of course, we write $\tilde{x} \geq \tilde{y}$ (or $\tilde{y} \leq \tilde{x}$) if the inequality holds componentwise; $\tilde{x} > \tilde{y}$ (or $\tilde{y} < \tilde{x}$) means that the strict inequality holds componentwise. Geometrically in $\mathbf{R}^2$, Figure 1 illustrates which vectors $\tilde{x}$ satisfy the two strict inequalities for a fixed vector $\tilde{y}$; the idea extends naturally to $\mathbf{R}^n$. We can also note that $\tilde{x} > \tilde{y}$ ($\tilde{x} \geq \tilde{y}$) means that $\tilde{x} - \tilde{y} \in \mathcal{O}$ ($\tilde{x} - \tilde{y} \in \bar{\mathcal{O}}$).

A fundamental theorem that is useful when considering monotonicity of solutions with respect to changes in an initial condition is the Kamke-Müller Theorem (presented with proof as Theorem 1.3.1 in [22]).

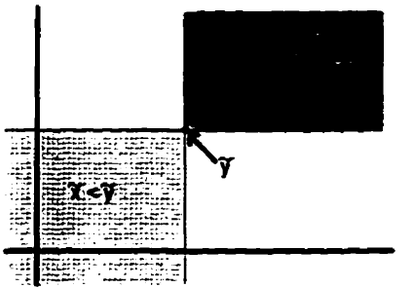


Figure 1: Orderings with respect to a quadrant in $\mathbf{R}^2$

Definition 2 *$\tilde{f}$ is* **quasimonotone non-decreasing** *if $\forall\, i$, $f_i(t,\tilde{u}) \leq f_i(t,\tilde{v})$ for $u_i = v_i$, $u_j \leq v_j$, $\forall\, j \neq i$.*

Theorem 1 (Kamke-Müller Theorem)

If $\dot{\tilde{v}} \leq \tilde{f}(t,\tilde{v})$, $\tilde{v}(0) = \tilde{v}_0$, and $\dot{\tilde{w}} \geq \tilde{f}(t,\tilde{w})$, $\tilde{w}(0) = \tilde{w}_0$, for $0 \leq t \leq T$, and $\tilde{v}_0 \leq \tilde{w}_0$, where $\tilde{f}(t,\tilde{x})$ is quasimonotone non-decreasing on $\Omega \subset \mathbf{R}^n$, Ω is open and convex, $\tilde{f}$ is continuous on Ω and Lipshitz continuous with respect to $\tilde{x}$ on compact subsets of Ω, then $\tilde{v}(t) \leq \tilde{w}(t)$ for $0 \leq t \leq T$.

Remark: The convexity assumption in Theorem 1 is not necessary. It is necessary if the quasimonotone non-decreasing hypothesis is replaced by

$$\frac{\partial f_i}{\partial x_j} \geq 0,\ i \neq j.$$

Suppose we are working with a chemical system and are interested in determining the monotonicity properties of the concentrations of species with respect to changes in initial concentrations. If the concentrations of the chemical species satisfy the autonomous system $\dot{\tilde{w}} = \tilde{f}(\tilde{w})$, $\tilde{w}(0) = \tilde{w}_0$, where $\tilde{f}$ is quasimonotone non-decreasing, then we can apply the Kamke-Müller Theorem to conclude that all concentrations are monotone non-decreasing with respect to changes in any initial condition. We may equivalently conclude that the partial derivative of any

concentration with respect to any initial concentration is non-negative. For either conclusion, we say that the system induces a *monotone flow* (with respect to the non-negative orthant).

Remark: A Mean Value Theorem argument shows that a solution component is monotone non-decreasing with respect to changes in an initial condition if and only if the partial derivative with respect to that initial condition of that component is non-negative for all time. Consider the system

$$\dot{\tilde{w}}(t) = \tilde{f}(t, \tilde{w}), \ \tilde{w}(0) = \tilde{w}_0. \tag{2.1}$$

Denote by $u_i(t, \tilde{u}_0)$ the i^{th} component of the solution to (2.1) with initial condition $\tilde{w}(0) = \tilde{u}(0)$. Denote by $u_i(t, \tilde{u}_0^*)$ the i^{th} component of the solution to (2.1) with initial condition $\tilde{w}(0) = \tilde{u}^*(0)$. Suppose that

$$u_j^*(0) > u_j(0), \text{ for some } j, \text{ and } u_i^*(0) = u_i(0), \ i \neq j. \tag{2.2}$$

If $u_i(t)$ is monotone non-decreasing with respect to changes in $u_j(0)$ – that is, if we know that $u_i(t, \tilde{u}_0^*) - u_i(t, \tilde{u}_0) \geq 0$ for all time and all pairs of vectors $\tilde{u}(0)$ and $\tilde{u}^*(0)$ satisfying (2.2), then we know that

$$\frac{\partial u_i}{\partial u_j(0)}(t) \geq 0$$

for all time.

By the Mean Value Theorem, for a fixed time t we know that there is some $\tilde{v}(0)$, with $u_j(0) \leq v_j(0) \leq u_j^*(0)$ and $u_i(0) = v_i(0)$, for $i \neq j$, such that

$$u_i(t, \tilde{u}_0^*) - u_i(t, \tilde{u}_0) = \{u_j^*(0) - u_j(0)\} \frac{\partial u_i}{\partial u_j(0)}(t, \tilde{v}_0);$$

hence, if we know that

$$\frac{\partial u_i}{\partial u_j(0)}(t) \geq 0$$

for all time, then $u_i(t)$ is monotone non-decreasing with respect to changes in $u_j(0)$. Furthermore, if we know that

$$\frac{\partial u_i}{\partial u_j(0)}(t) > 0$$

for all time, then $u_i(t)$ is strictly monotone increasing with respect to changes in $u_j(0)$.

For non-monotonicity, we can show by a Taylor series argument that if the partial derivative with respect to an initial value of a solution component takes both signs in time, then two solution curves with that initial condition different but sufficiently close will cross. The Taylor series in $u_j(0)$ for $u_i(t, u_0)$ near $u_j^*(0)$ is given by

$$u_i(t, \tilde{u}_0^*) - u_i(t, \tilde{u}_0) = (u_j^*(0) - u_j(0))\frac{\partial u_i}{\partial u_j(0)}(t, \tilde{u}_0) + O((u_j^*(0) - u_j(0))^2).$$

So, if $\frac{\partial u_i}{\partial u_j(0)}(t, \tilde{u}_0)$ is of both signs in time, then for $u_j^*(0)$ sufficiently close to $u_j(0)$ we can conclude that the curves $u_i(t, \tilde{u}_0)$ and $u_i(t, \tilde{u}_0^*)$ cross. This ends the remark.

Recall that a flow ϕ for an autonomous system $\dot{\tilde{x}}(t) = \tilde{f}(\tilde{x})$ is defined to be the map $\phi_t : \Omega \mapsto \Omega$ where $\phi_t(\tilde{x})$ is the solution with initial value $\tilde{x}$. We define the positive interval of existence $I(\tilde{x})$ by

$$I(\tilde{x}) = \{t \geq 0 : \phi_t(\tilde{x}) \in \Omega\}.$$

If $t \in I(\tilde{x})$ and $s \in I(\phi_t(\tilde{x}))$, then

$$\begin{aligned} s + t &\in I(\tilde{x}) \\ \phi_s(\phi_t(\tilde{x})) &= \phi_{s+t}(\tilde{x}), \text{ and} \\ \phi_0(\tilde{x}) &= \tilde{x}, \ \forall\, \tilde{x} \in \Omega. \end{aligned}$$

Definition 3 *The flow ϕ is* **monotone** *if*

$$\tilde{w}_0 \leq \tilde{v}_0 \ \Rightarrow \ \phi_t(\tilde{w}_0) \leq \phi_t(\tilde{v}_0), \ \forall\, t \in I(\tilde{w}) \cap I(\tilde{v}). \tag{2.3}$$

Reference [12] offers a pleasant discussion of monotone flows.

Example 1: All monomolecular chemical reactions induce a monotone flow. Let A_i, $i = 1, \ldots, n$, denote n reacting substances and let $x_i(t)$, $i = 1, \ldots, n$, denote the concentration of A_i at time t. In a monomolecular reaction mechanism, all the reactions are of the form

$$A_j \xrightarrow{k} A_l \tag{2.4}$$

which means that all the rate expressions are linear. Thus we obtain a linear, constant coefficient system of differential equations,

$$\dot{\tilde{x}} = \mathcal{A}\tilde{x}(t), \ \tilde{x}(0) = \tilde{x}_0, \tag{2.5}$$

where the matrix $\mathcal{A} = (a_{ij})$ satisfies:

(1) $a_{ii} \leq 0 \ \forall i$,

(2) $a_{ij} \geq 0 \ \forall \ i \neq j$,

(3) $\sum_{i=1}^{n} a_{ij} = 0$, and

(4) No row of $\mathcal{A}$ is $\tilde{0}$.

All these properties follow directly from the form of the rate expressions and the induced DEs (see [30]). Solving (2.5),

$$\begin{aligned} \tilde{x}(t) &= \tilde{x}_0 e^{\mathcal{A}t} \\ &= \tilde{x}_0 e^{-\lambda I t} e^{(\mathcal{A}+\lambda I)t} \\ &= \tilde{x}_0 e^{-\lambda t} e^{(\mathcal{A}+\lambda I)t} \end{aligned} \tag{2.6}$$

where λ is some positive number. For λ sufficiently large, all entries in the matrix $\mathcal{A} + \lambda I$ are non-negative. Hence, all entries in

$$\mathcal{B}(t) = (b_{ij}(t)) = e^{(\mathcal{A}+\lambda I)t}$$

are non-negative, $t \geq 0$. Now, for any i,

$$x_i(t) = e^{-\lambda t} \sum_{i=1}^{n} b_{ij}(t) x_j(0),$$

and since $b_{ij}(t) \geq 0, \forall i, j$, we get

$$\frac{\partial x_i}{\partial (x_m(0))}(t) = b_{im}(t) e^{-\lambda t} \geq 0, \forall\, t \geq 0, \forall\, i, m;$$

all concentrations are monotone with respect to changes in any initial concentrations.

In [34], it is shown that a necessary and sufficient condition for the autonomous system

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}),\ \tilde{x}(0) = \tilde{x}_0, \tag{2.7}$$

to induce a monotone flow is

$$f_{i,j} = \frac{\partial f_i}{\partial x_j}(\tilde{x}) \geq 0,\ \forall\, i \neq j. \tag{2.8}$$

M. Hirsch discusses cooperative or competitive vector fields in [11] and [13]. Here, "cooperative" means that $f_{i,j} \geq 0$, for $i \neq j$, and "competitive" means that $f_{i,j} \leq 0$, for $i \neq j$. A new generalization of a result in these papers will be discussed later.

When we do not have a monotone flow, we can consider the possibility of there being a simple transformation to the system, which essentially switches the signs of some of the components of the system, in order to produce a monotone flow. In this case, the original system is called an *order preserving flow* and all solution components are still monotone with respect to changes in any solution component. Mathematically, we say that this system induces an order preserving flow with respect to an orthant if there is a matrix P,

$$P = diag[(-1)^{m_1}, \ldots, (-1)^{m_n}],\ m_i \in \{0, 1\}, \tag{2.9}$$

such that under the coordinate transformation $\tilde{y} = P\tilde{x}$, we obtain a monotone flow in $\tilde{y}$; in other words,

$$P\tilde{w}_0 \leq P\tilde{v}_0 \Rightarrow P\phi_t(\tilde{w}_0) \leq P\phi_t(\tilde{v}_0), \forall\, t \in I(\tilde{w}_0) \cap I(\tilde{v}_0).$$

Lemma 2.1 in [34] gives a necessary and sufficient condition for an order preserving flow, namely that

$$\text{all off-diagonal entries of } PD\tilde{f}(\tilde{x})P \text{ are non-negative,} \tag{2.10}$$

where $D\tilde{f}(\tilde{x})$ represents the $n \times n$ Jacobian matrix with $f_{i,j}$ as its $(i,j)^{th}$ entry. Some work from [18] appears in [20]. It was shown that for every possible order preserving flow sign pattern a chemical reaction mechanism which induces it can be constructed and, in fact, if the mechanism is to induce an order preserving flow, only certain chemical reactions are allowed.

2.2 Practical Tools for Establishing Monotonicity

Unfortunately, most systems (2.7) do not satisfy the restrictive conditions for monotone or order preserving flows. In [18], the monotonicity properties of several chemical and epidemiological models were investigated. None of the considered problems induced monotone or order preserving flows; so specific, and somewhat unsophisticated, methods were used to analyze each problem. More refined techniques have since been developed.

To give the reader an understanding of the thought process followed in establishing the upcoming results, we very quickly mention the theory of qualitative stability. The question of qualitative (or sign) stability of the system

$$\dot{\tilde{x}}(t) = \mathcal{A}\tilde{x}, \tag{2.11}$$

where $\mathcal{A} = (a_{ij})$ is a constant matrix, has been solved. The system is said to be stable if all of the eigenvalues of $\mathcal{A}$ have negative real parts. The system is said to be sign stable if for each matrix $\mathcal{B} = (b_{ij})$ with $sign(b_{ij}) = sign(a_{ij})$, $\forall\ i, j$, the corresponding system, $\dot{\tilde{x}}(t) = \mathcal{B}\tilde{x}$, is stable regardless of the magnitudes of the elements b_{ij}. A graph theoretic approach proved to be essential in solving this problem (see [26], [9], [15], and [17]; the interested reader is warned that results were restated incorrectly in the literature). Upon examination of the theory of sign stability, it seems natural to consider a graph theoretic approach to the problem of monotonicity with respect to changes in an initial condition of the nonlinear system of differential equations (2.7).

2.2.1 A Graph Theoretic Approach

Subsections 2.2.1 through 2.2.4 present the work contained in [19] and [21].

We associate with the matrix $D\tilde{f}$ a signed, directed multigraph. For $S \subset \Omega$, let $G(\tilde{f}, S)$ be the signed, directed multigraph with vertices labelled $v_1, \ldots, v_n$, where vertex v_i is associated with solution component x_i, constructed in the following way:

G.i) If $f_{i,j} > 0$, $i \neq j$, at some point of S, a positive edge, labelled e^+_{ji}, directed from vertex v_j to vertex v_i is drawn in the multigraph.

G.ii) If $f_{i,j} < 0$, $i \neq j$, at some point of S, a negative edge, labelled e^-_{ji}, directed from vertex v_j to vertex v_i is drawn in the multigraph.

G.iii) If $f_{i,j} = 0$, $\forall\ \tilde{x} \in S$, $i \neq j$, no edge is drawn in the graph.

Note: if $f_{i,j}$ takes both signs in S, (G.i) and (G.ii) will both apply.

Edges in the directed multigraph are *parallel* if they have the same end vertices and direction; edges are anti-parallel if they have the same end vertices and opposite directions. Duplicate edges (same sign and direction) of $G(\tilde{f}, S)$ may be deleted. It should be noted that there are no directed edges from a vertex to itself. Two distinct vertices in $G(\tilde{f}, S)$ can be connected by at most two parallel and two anti-parallel directed edges, one edge of each sign in each direction. This leads to a potential difficulty in labelling edges, which we deal with by including the sign in the edge label. An arbitrary edge of $G(\tilde{f}, S)$ will be denoted e_{ij}^{s}, $s \in \{+, -\}$. In the examples, positive edges will be drawn as solid lines and negative edges will be drawn as dashed lines. Note that $G(\tilde{f}, S)$ need not be connected. In the examples, it will be.

At times, we will assume that

$H(\tilde{f}, S)$ $\quad f_{i,j}(\tilde{x}) \geq 0$ or $f_{i,j}(\tilde{x}) \leq 0, \forall\, \tilde{x} \in S, \forall\, i \neq j.$

Under $H(\tilde{f}, S)$, when $D\tilde{f}$ is sign symmetric, $sign(f_{i,j}(\tilde{x}))sign(f_{j,i}(\tilde{x})) \geq 0$, $\forall\, \tilde{x} \in S$, $\forall\, i \neq j$, and any parallel or anti-parallel edges in $G(\tilde{f}, S)$ will have the same sign. We construct the signed (undirected) multigraph $G_u(\tilde{f}, S)$ by removing the directions of edges in $G(\tilde{f}, S)$ and deleting any redundant parallel edges. In an undirected graph, edges are parallel if they have the same end vertices; "redundant parallel edges" are parallel edges with the same sign. In [20], $H(\tilde{f}, S)$ is not introduced and a slightly different $G_u(\tilde{f}, S)$, allowing parallel edges of opposite sign, is used. [20] shows how to apply this graph theory directly to chemical mechanisms, without listing the induced system of differential equations, making hypothesis $H(\tilde{f}, S)$ unnatural to implement.

2.2.2 Notation and Terminology

The following concepts will be required in the upcoming graphical discussion.

A *walk* (*path*) in $G_u(\tilde{f}, S)$ between two vertices, v_i and v_j, is an alternating sequence of vertices and edges (distinct vertices and distinct edges) beginning with v_i and ending with v_j. A *directed walk* (*directed path*) in $G(\tilde{f}, S)$ from vertex v_i to vertex v_j is an alternating sequence of vertices and directed edges (distinct vertices and distinct directed edges) beginning with v_i and ending with v_j, with the edges being appropriately directed. A directed path (directed walk) can be described by a sequence of vertices and edges (distinct vertices and distinct directed edges) as $v_i e_{ik_1}^{s_0} v_{k_1} e_{k_1k_2}^{s_1} v_{k_2} \ldots e_{k_mj}^{s_m} v_j$, where edge $e_{r_1r_2}^{s_r}$ has tail vertex v_{r_1} and head vertex v_{r_2} and $s_r \in \{+, -\}$. In a *closed walk*, $v_j = v_i$.

The *length* of a walk (path) is the total number of edges comprising it.

The *sign* of a walk (path) is the product of the signs of the edges comprising it.

Combining a path and an edge which connects the terminal vertices of the path creates a *cycle* in $G_u(\tilde{f}, S)$. A *directed cycle* in $G(\tilde{f}, S)$ is a directed walk $v_i e_{ik_1}^{s_0} v_{k_1} e_{k_1k_2}^{s_1} v_{k_2} \ldots v_{k_{m-1}} e_{k_{m-1}k_m}^{s_m} v_{k_m}$, where $v_i e_{ik_1}^{s_0} v_{k_1} e_{k_1k_2}^{s_1} v_{k_2} \ldots v_{k_{m-1}}$ is a directed path, $v_i = v_{k_m}$, and $s_r \in \{+, -\}$.

The vertex v_i is a *source* (*sink*) if all incident edges are outgoing (incoming).

We say that the ordered vertex pair (v_i, v_j), $i \neq j$, is *strongly connected* if there is a directed path from v_i to v_j. Notice that if there is a directed walk from v_i to v_j then there is a directed path from v_i to v_j as well. The directed graph is *strongly connected* if for each ordered pair of vertices (v_i, v_j) in the graph, (v_i, v_j) is strongly connected. Furthermore, a strongly connected vertex pair (v_i, v_j), $i \neq j$, is

(i) *positively* (*negatively*) *consistently strongly connected* if all directed walks

from v_i to v_j are positive (negative), and

(ii) *inconsistently strongly connected* if there is a directed walk of each sign from v_i to v_j.

Notice that we must use "walk" above; should the walk involve a negative directed cycle, then the ordered vertex pair will be inconsistently strongly connected.

We say (v_i, v_i) is *inconsistently strongly connected* if v_i is part of a negative directed cycle. Otherwise, we say that (v_i, v_i) is *positively consistently strongly connected.* If (v_i, v_j) is not inconsistently strongly connected, then we say that (v_i, v_j) is *consistent.*

Given a connected graph, a *spanning tree* of this graph is a connected, spanning subgraph without any cycles. A *spanning subgraph* is a subgraph involving all of the vertices of the original graph. An unconnected graph has a *spanning forest*, a collection of spanning trees, one for each of its connected components. Adding back an edge from the original graph which is not in a spanning forest produces a subgraph with exactly one cycle. Separately adding each of the excluded edges to a spanning forest gives a set of cycles called a *fundamental set of cycles.* Let F be a spanning forest in a graph G. For each edge e of $G-F$, there is a unique cycle $\mathcal{C}_e$ such that $\mathcal{C}_e - F = \{e\}$. For any F, $\{\mathcal{C}_e : e \in G-F\}$ is a fundamental set of cycles. A fundamental set of cycles forms a basis for the set of all unions of edge-disjoint cycles in the graph; here, we are thinking of a basis in the linear algebra sense where the field is $\mathbb{Z}_2$. Let $\mathcal{C}_1 = \{e_1, \ldots, e_p, e_{p+1}, \ldots, e_q\}$ and $\mathcal{C}_2 = \{e_1, \ldots, e_p, e_{q+1}, \ldots, e_r\}$ be two unions of edge-disjoint cycles. Then the sum is $\mathcal{C}_1 + \mathcal{C}_2 = \{e_{p+1}, \ldots, e_r\}$. This is the symmetric difference of the sets, namely $\mathcal{C}_1 + \mathcal{C}_2 = (\mathcal{C}_1 - \mathcal{C}_2) \cup (\mathcal{C}_2 - \mathcal{C}_1)$. We present the following simple lemma.

Lemma 2 *For C_1 and C_2 two unions of edge-disjoint cycles,*

$$sign(C_1 + C_2) = sign(C_1)sign(C_2).$$

Proof: Suppose $C_1 = \{e_1, \ldots, e_p, e_{p+1}, \ldots, e_q\}$ and $C_2 = \{e_1, \ldots, e_p, e_{q+1}, \ldots, e_r\}$ are two unions of edge disjoint cycles. Then $C_1 + C_2 = \{e_{p+1}, \ldots, e_r\}$ and

$$\begin{aligned} sign(C_1 + C_2) &= \prod_{i=p+1}^{r} sign(e_i) = \prod_{i=p+1}^{r} sign(e_i) \left(\prod_{i=1}^{p} sign(e_i) \right)^2 \\ &= sign(C_1)sign(C_2). \end{aligned}$$

□

Remark: Some treatments of elementary graph theory define cycle to include unions of edge-disjoint cycles. This permits the above algebra to be cleanly stated: the space of cycles (so defined) is closed under *mod* 2 addition. In this work, we find the simpler definition of cycle to be more useful.

2.2.3 Graph Theoretic Results

These first results deal with system-wide monotonicity. Theorem 5 gives a graph theoretical equivalent to condition (2.10) for an order preserving flow. We will need these simple observations before presenting this result.

Lemma 3 *Every closed, negative walk contains a negative cycle.*

Proof: The proof is by induction on the length of the closed, negative walk. Suppose the walk has length 2; then it must consist of one edge of each sign and the result follows. Suppose the result is true for a closed, negative walk of length $\leq m$. Consider a closed, negative walk of length $m + 1$. If the closed, negative walk has no repeated vertices, then the result follows. If the closed, negative walk has a

repeated vertex, then the walk is a union of two closed walks, one of each sign. The closed, negative walk in this union has length $\leq m-1$ and, therefore, must contain a negative cycle by the induction hypothesis. □

Lemma 4 *If $\tilde{f}$ induces an order preserving flow and the graph $G(\tilde{f},\Omega)$ has a positive (negative) directed edge labelled e_{ij}^{+} (e_{ij}^{-}) from vertex v_i to vertex v_j, $i \neq j$, then $P_iP_j = 1$ (-1) where $P = diag[P_i]$ is the matrix associated with the order preserving flow.*

Proof: Suppose the directed edge has positive sign. Then $f_{j,i} \geq 0$ in Ω and $f_{j,i} > 0$ at some point of Ω. Since $\tilde{f}$ gives an order preserving flow, there is a matrix P as in (2.9) such that $P_jf_{j,i}P_i \geq 0$. Evaluating at the point where $f_{j,i} > 0$, we must have $P_jP_i > 0$ and hence $P_iP_j = 1$. The other case is argued in the same way. □

Theorem 5 *System (2.7) induces an order preserving flow if and only if the conditions $H(\tilde{f},\Omega)$ and $sign(f_{i,j}(\tilde{x}))sign(f_{j,i}(\tilde{x})) \geq 0$, $\forall \tilde{x} \in \Omega$, $\forall\, i \neq j$, hold and there are either no cycles in $G_u(\tilde{f},\Omega)$ or every cycle in any one fundamental set of cycles in $G_u(\tilde{f},\Omega)$ is positive.*

Proof: We first establish that (2.7) induces an order preserving flow if and only if condition $H(\tilde{f},\Omega)$ holds and $G_u(\tilde{f},\Omega)$ contains no negative cycles. The forward direction of the Theorem follows immediately; the other direction follows from Lemma 2.

If $\tilde{f}$ induces an order preserving flow, there is a matrix P as in (2.9). Suppose that the graph $G_u(\tilde{f},\Omega)$ has a negative cycle with vertices $v_{l_1},\ldots,v_{l_k}$, where v_{l_i} is connected to $v_{l_{i+1}}$, $1 \leq i \leq k$, with the convention $v_{l_{k+1}} = v_{l_1}$. By Lemma 4, $P_{l_i}P_{l_{i+1}}$ is the sign of the edge between v_{l_i} and $v_{l_{i+1}}$. Since the cycle is negative

we must have $(P_{l_1}P_{l_2})(P_{l_2}P_{l_3})\cdots(P_{l_k}P_{l_1}) = -1$, but this is a contradiction because $(P_{l_1}P_{l_2})(P_{l_2}P_{l_3})\cdots(P_{l_k}P_{l_1}) = (P_{l_1}P_{l_2}\cdots P_{l_k})^2 = 1$.

Next, we prove that if $G_u(\tilde{f},\Omega)$ has no negative cycles then we must have an order preserving flow. The graph $G_u(\tilde{f},\Omega)$ consists of connected components $G_u^1(\tilde{f},\Omega),\ldots,G_u^p(\tilde{f},\Omega)$. Since the variables corresponding to the vertices in two different subgraphs do not interact, we need only consider a connected subgraph of $G_u(\tilde{f},\Omega)$, say $G_u^1(\tilde{f},\Omega)$, with n_1 vertices.

Choose any vertex v_1 in $G_u^1(\tilde{f},\Omega)$. Since there are no negative cycles, every vertex in the subgraph is connected to v_1 by paths of only one sign. If not, then v_1 is connected to v_k, say, by paths of both sign; hence, v_1 is part of a closed, negative trail (combining the two paths) and, by Lemma 3, $G_u^1(\tilde{f},\Omega)$ contains a negative cycle, giving a contradiction. We define the disjoint sets

$$\begin{aligned}\mathcal{Q} &= \{v_k : v_1 \text{ and } v_k \text{ are only connected by positive paths in } G_u^1(\tilde{f},\Omega)\}, \text{ and}\\ \mathcal{R} &= \{v_k : v_1 \text{ and } v_k \text{ are only connected by negative paths in } G_u^1(\tilde{f},\Omega)\}.\end{aligned}$$

In order to avoid a simple contradiction, vertices in $\mathcal{Q}$ can only be connected to each other by positive paths, vertices in $\mathcal{R}$ can only be connected to each other by positive paths, and vertices in $\mathcal{Q}$ can only be connected to vertices in $\mathcal{R}$ by negative paths. We relabel the vertices in $G_u^1(\tilde{f},\Omega)$ so that $v_2,\ldots,v_q \in \mathcal{Q}$ and $v_{q+1},\ldots,v_{n_1} \in \mathcal{R}$. Hence,

$$\begin{aligned}&f_{i,k} \geq 0, \quad 1 \leq i,k \leq q,\ i \neq k,\\ &f_{i,k} \geq 0, \quad q+1 \leq i,k \leq n_1,\ i \neq k,\\ &f_{i,k} \leq 0, \quad 1 \leq i \leq q,\ \ q+1 \leq k \leq n_1,\\ &f_{i,k} \leq 0, \quad 1 \leq k \leq q,\ \ q+1 \leq i \leq n_1.\end{aligned}$$

For this subsystem, $D\tilde{f}$ has the sign pattern given in Figure 2, where '+' means the

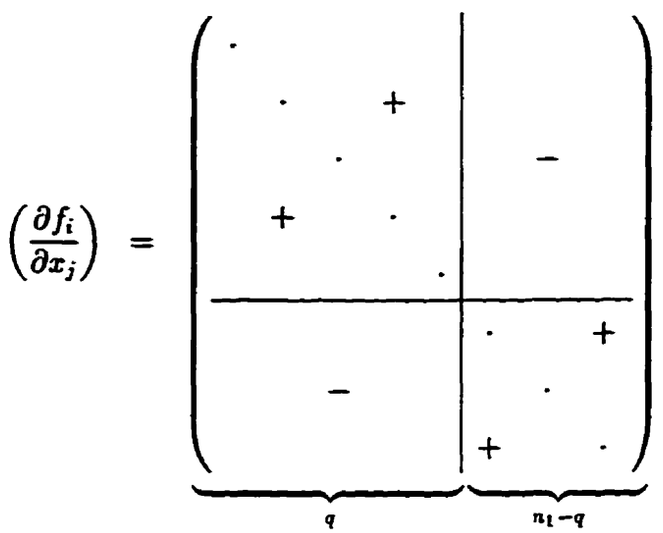


Figure 2: Sign pattern of $D\tilde{f}$ after relabelling.

corresponding partial derivative is non-negative and '$-$' means it is non-positive; hence, choosing the matrix $P = diag[1, \ldots, 1, -1, \ldots, -1]$, where P has q entries of 1 and $n_1 - q$ entries of -1, means that $P(D\tilde{f})P$ has non-negative off-diagonal entries. Thus, this subsystem gives an order preserving flow. The result follows. □

A version of Theorem 5 was stated informally in [34] and, later, more carefully in [35]; however, a careful proof has not been found in the literature.

Remark: Let $\mathcal{C}$ be a fundamental set of cycles in a graph. Every cycle in the graph is positive if and only if every cycle in $\mathcal{C}$ is positive; thus, in checking for an order preserving flow, only a single fundamental set of cycles needs to be considered. The number of cycles in a fundamental set of cycles is bounded by the number of edges in the graph. The total number of cycles can be exponential in the number of edges.

We may want to apply this result to the non-negative orthant. In this case, we will suppose that $\tilde{f} \in C^1(\bar{\Omega})$. The following limit property is necessary (see [35]);

it extends our ordering to the non-negative orthant.

Proposition 6 *Let Ω be an open, connected, convex set. Then for any $\tilde{x}^{(*)}, \tilde{y}^{(*)} \in \bar{\Omega}$ with $P\tilde{x}^{(*)} \leq P\tilde{y}^{(*)}$, $\exists$ sequences $\{\tilde{x}^{(m)}\}, \{\tilde{y}^{(m)}\} \in \Omega$ with $P\tilde{x}^{(m)} \leq P\tilde{y}^{(m)}, \forall\, m$, and*

$$\lim_{m\to\infty} \tilde{x}^{(m)} = \tilde{x}^{(*)} \text{ and } \lim_{m\to\infty} \tilde{y}^{(m)} = \tilde{y}^{(*)}.$$

Proof: Let $\tilde{\eta} \in \Omega$. For $\tilde{x}^{(*)}$ and $\tilde{y}^{(*)} \in \bar{\Omega}$, the sequences $\tilde{x}^{(m)}$ and $\tilde{y}^{(m)}$ in Ω can be defined by $\tilde{x}^{(m)} = \tilde{x}^{(*)} + 2^{-m}(\tilde{\eta} - \tilde{x}^{(*)})$ and $\tilde{y}^{(m)} = \tilde{y}^{(*)} + 2^{-m}(\tilde{\eta} - \tilde{y}^{(*)})$; then $P(\tilde{x}^{(m)} - \tilde{y}^{(m)}) = (1 - 2^{-m})P(\tilde{x}^{(*)} - \tilde{y}^{(*)})$ and the limit property holds. □

Proposition 6 implies that if $P\tilde{x} \leq P\tilde{y} \Rightarrow P\varphi_t(\tilde{x}) \leq P\varphi_t(\tilde{y})$ for $\tilde{x}, \tilde{y} \in \Omega$, then $P\tilde{x}^{(*)} \leq P\tilde{y}^{(*)} \Rightarrow P\varphi_t(\tilde{x}^{(*)}) \leq P\varphi_t(\tilde{y}^{(*)})$. This leads us to the following result.

Corollary 7 *Suppose $\bar{\Omega}$ is a closed, positively invariant, connected, convex set. System (2.7) induces an order preserving flow with respect to an orthant if and only if the conditions $H(\tilde{f}, \Omega)$ and $sign(f_{i,j}(\tilde{x}))sign(f_{j,i}(\tilde{x})) \geq 0$, $\forall\ \tilde{x} \in \Omega$, $\forall\ i \neq j$, hold and there are either no cycles in the graph of $D\tilde{f}$ or every cycle in any one fundamental set of cycles in the graph of $D\tilde{f}$ is positive.*

As we will be focusing from now on on the signs of partial derivatives of components with respect to initial conditions we can drop the convexity assumption on Ω. Furthermore, it will suffice to make assumptions only on $G(\tilde{f}, \Gamma(\tilde{x}))$ rather than $G(\tilde{f}, \Omega)$, where $\Gamma(\tilde{x}) = \{\phi_t(\tilde{x}) : t \geq 0\}$ is the positive semi-trajectory. The following theorem shows how to determine the signs of partial derivatives with respect to initial conditions when $G(\tilde{f}, \Gamma(\tilde{x}))$ has no negative cycles.

Theorem 8 *Suppose that $G_u(\tilde{f}, \Gamma(\tilde{x}))$ has no negative cycles. If all paths connecting v_i and v_j in $G_u(\tilde{f}, \Gamma(\tilde{x}))$ are positive then*

$$\frac{\partial \phi_t^j}{\partial x_i}(\tilde{x}) \geq 0 \text{ and } \frac{\partial \phi_t^i}{\partial x_j}(\tilde{x}) \geq 0,$$

$\forall\ t \in I(\tilde{x})$. *If all paths connecting v_i and v_j in $G_u(\tilde{f}, \Gamma(\tilde{x}))$ are negative, then the derivatives are both non-positive.*

Proof: If all paths connecting v_i and v_j in $G_u(\tilde{f}, \Gamma(\tilde{x}))$ are of one sign then v_i and v_j must be part of a connected subgraph $G_u^1(\tilde{f}, \Gamma(\tilde{x}))$ which has no negative cycles. The argument used to prove Theorem 5 then applies and gives the result. □

These final theorems allow us to obtain partial and strong monotonicity results; they require considering the graph along the solution curve.

Theorem 9 *If the vertex pair (v_i, v_i) is positively consistently strongly connected in $G(\tilde{f}, \Gamma(\tilde{x}))$, then*

$$\frac{\partial \varphi_t^i}{\partial x_i}(\tilde{x}) > 0,\ \forall\ t \in I(\tilde{x}).$$

Furthermore, if (v_i, v_j), $i \neq j$, is consistent in $G(\tilde{f}, \Gamma(\tilde{x}))$ then, $\forall\ t \in I(\tilde{x})$,

$$\frac{\partial \varphi_t^j}{\partial x_i}(\tilde{x}) \begin{cases} = 0 \text{ if } (v_i, v_j) \text{ is not strongly connected in } G(\tilde{f}, \Gamma(\tilde{x})) \\ \geq 0 \text{ if } (v_i, v_j) \text{ is positively consistently strongly connected} \\ \quad \text{in } G(\tilde{f}, \Gamma(\tilde{x})) \\ \leq 0 \text{ if } (v_i, v_j) \text{ is negatively consistently strongly connected} \\ \quad \text{in } G(\tilde{f}, \Gamma(\tilde{x})) \end{cases}$$

Proof: Suppose (v_1, v_j) is consistent in $G(\tilde{f}, \Gamma(\tilde{x}))$. For any vertex v_k, (v_1, v_k) is either consistently strongly connected, inconsistently strongly connected or not

strongly connected. Define the disjoint sets

$$\begin{aligned}
\mathcal{Q}_1 &= \{v_k : (v_1, v_k) \text{ is positively consistently strongly connected in } G(\tilde{f}, \Gamma(\tilde{x}))\},\\
\mathcal{Q}_2 &= \{v_k : (v_1, v_k) \text{ is negatively consistently strongly connected in } G(\tilde{f}, \Gamma(\tilde{x}))\},\\
\mathcal{R} &= \{v_k : (v_1, v_k) \text{ is not strongly connected in } G(\tilde{f}, \Gamma(\tilde{x}))\}, \text{ and}\\
\mathcal{S} &= \{v_k : (v_1, v_k) \text{ is inconsistently strongly connected in } G(\tilde{f}, \Gamma(\tilde{x}))\}.
\end{aligned}$$

We relabel the vertices so that $v_1, \ldots, v_{q_1} \in \mathcal{Q}_1$, $v_{q_1+1}, \ldots, v_{q_2} \in \mathcal{Q}_2$, $v_{q_2+1}, \ldots, v_r \in \mathcal{R}$, and $v_{r+1}, \ldots, v_n \in \mathcal{S}$. With the corresponding relabelling of x_i, $1 \le i \le n$, the system (2.7) takes on the form

$$\begin{aligned}
\dot{x}_i &= f_i(x_1, \ldots, x_{q_2}, x_{q_2+1}, \ldots, x_r), && 1 \le i \le q_2,\\
\dot{x}_i &= f_i(x_{q_2+1}, \ldots, x_r), && q_2 + 1 \le i \le r,\\
\dot{x}_i &= f_i(x_1, \ldots, x_{q_2}, x_{q_2+1}, \ldots, x_r, x_{r+1}, \ldots, x_n), && r + 1 \le i \le n.
\end{aligned}$$

Since (v_1, v_j) is consistent, either

(i) $q_2 + 1 \le j \le r$, in which case

$$\frac{\partial \varphi_t^j}{\partial x_1}(\tilde{x}) = 0, \ \forall\, t \ge 0,$$

or

(ii) $1 \le j \le q_2$, and then

$$\begin{aligned}
\frac{\partial \dot{\varphi}_t^j}{\partial x_1}(\tilde{x}) &= \sum_{k=1}^{q_2} f_{j,k} \frac{\partial \varphi_t^k}{\partial x_1}(\tilde{x}) + \sum_{k=q_2+1}^{r} f_{,k} \frac{\partial \varphi_t^k}{\partial x_1}(\tilde{x})\\
&= \sum_{k=1}^{q_2} f_{j,k} \frac{\partial \varphi_t^k}{\partial x_1}(\tilde{x}). \qquad (2.12)
\end{aligned}$$

We can write (2.12) as $\dot{\tilde{Y}}(t) = M\tilde{Y}(t)$ where $\tilde{Y}$ is a q_2 vector and M is a $q_2 \times q_2$ matrix with $Y_j = \frac{\partial \varphi_t^j}{\partial x_1}$ and $M_{jk} = f_{j,k}$. Choosing

$$P = diag[P_i] = diag[1, \ldots, 1, -1, \ldots, -1], \qquad (2.13)$$

where P has q_1 entries of 1 and $q_2 - q_1$ entries of -1, means that PMP has non-negative off-diagonal entries.

We let $\tilde{Z} = P\tilde{Y}$. Since $\tilde{Y}(0) = \tilde{E}$, then $\tilde{Z}(0) = \tilde{E}$, where $\tilde{E} = (1, 0, \ldots, 0)^T$. We will prove that $\tilde{Z} \geq \tilde{0}$, $\forall\ t \in I(\tilde{x})$, which by the construction of $\tilde{Z}$ gives the remainder of the second result (with $i = 1$). Now,

$$\dot{\tilde{Y}} = P\dot{\tilde{Z}} = M\tilde{Y} = MP\tilde{Z}.$$

So, we have

$$\dot{\tilde{Z}} = (PMP)\tilde{Z}. \tag{2.14}$$

Since $(PMP)\tilde{Z}$ is quasimonotone nondecreasing, $\tilde{Z}(t) \geq \tilde{0}$, $\forall\ t \geq 0$, by the Kamke-Müller Theorem.

We still need to prove the first result. Take $j = 1$ in the above setup. We will show that $Z_1(t) > 0, \forall\ t \in I(\tilde{x})$. From (2.14), we have

$$\dot{\tilde{Z}} + \lambda\tilde{Z} = (PMP + \lambda)\tilde{Z} = N\tilde{Z},$$

where λ is chosen large enough so that $N = PMP + \lambda I \geq 0$. Solving for $\tilde{Z}$ in terms of the right-hand side, we get

$$\tilde{Z}(t) = e^{\lambda t}\tilde{E} + \int_0^t e^{\lambda(s-t)} N\tilde{Z}(s)ds. \tag{2.15}$$

We can immediately conclude from (2.15) that $Z_1(t) > 0$, $\forall\ t \in I(\tilde{x})$. □

Definition 4 *We say that (2.7) gives a* consistent flow *if, for each i and j, either*

$$\frac{\partial\varphi_t^j}{\partial x_i}(\tilde{x}) \geq 0,\ \forall\ t \in I(\tilde{x}),\ \textit{or}\ \frac{\partial\varphi_t^j}{\partial x_i}(\tilde{x}) \leq 0,\ \forall\ t \in I(\tilde{x}).$$

If (v_i, v_j) is consistent in $G(\tilde{f}, S)$ for each i and j, we say that $G(\tilde{f}, S)$ is consistent.

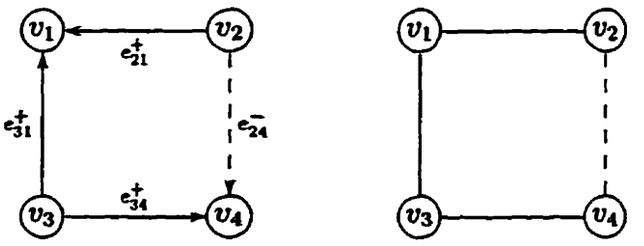


Figure 3: $G(\tilde{f}, \Omega)$ and $G_u(\tilde{f}, \Omega)$ for Example 2.

Thus, by Theorem 8, we get the following corollary.

Corollary 10 *If (2.7) gives a consistent graph $G(\tilde{f}, \Omega)$, then (2.7) gives a consistent flow.*

Example 2: Consider the example graphs $G(\tilde{f}, \Omega)$ and $G_u(\tilde{f}, \Omega)$ in Figure 3. As mentioned earlier, we adopt the convention of positive edges being represented by solid lines and negative edges being represented by dashed lines. $G_u(\tilde{f}, \Omega)$ consists of a single negative cycle and, by Theorem 5, does not induce an order preserving flow. However, there are no inconsistently strongly connected vertices. This is a consistent flow and we can immediately state the sign pattern for the matrix of partial derivatives of solution components with respect to initial conditions, namely

$$\left(\frac{\partial \varphi_t^i}{\partial x_j}(\tilde{x})\right) = \begin{pmatrix} + & 0 & 0 & 0 \\ + & + & 0 & - \\ + & 0 & + & + \\ 0 & 0 & 0 & + \end{pmatrix},$$

where '+' means that the corresponding partial derivative is non-negative, '−' means that the corresponding partial derivative is non-positive, and '0' means that the corresponding partial derivative is identically zero.

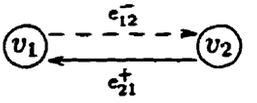
We can also use Theorem 8 to give partial results on the signs of partial derivatives with respect to initial values.

The condition that $G(\tilde{f}, \Omega)$ be consistent is a sufficient, but not a necessary, condition for a consistent flow as the following example illustrates.

Example 3: Consider

$$\dot{x}_1 = f_1 = \begin{cases} x_2^2, & x_2 \geq 0 \\ 0, & x_2 < 0 \end{cases}, \quad \dot{x}_2 = f_2 = \begin{cases} 0, & x_2 \geq 0 \\ -x_1 x_2^2, & x_2 < 0 \end{cases}.$$

One can check that $f_{1,2} \geq 0$ and $f_{2,1} \leq 0$, $\forall\, \tilde{x} \in \Omega = \mathbb{R}^2$, so the graph, $G(\tilde{f}, \Omega)$, for this system is as follows:



The vertices are inconsistently strongly connected and Theorem 8 does not apply. Yet, solving the system gives:

$$\varphi_t^1(x_1, x_2) = \begin{cases} (x_2)^2 t + x_1, & x_2 \geq 0 \\ x_1, & x_2 < 0 \end{cases}, \quad \varphi_t^2(x_1, x_2) = \begin{cases} x_2, & x_2 \geq 0 \\ \frac{x_2}{1 + x_1 x_2 t}, & x_2 < 0 \end{cases}.$$

This gives the following sign pattern for partial derivatives with respect to initial conditions:

$$\left(\frac{\partial \varphi_t^i}{\partial x_j}(\tilde{x})\right) = \begin{pmatrix} + & + \\ - & + \end{pmatrix}$$

where '+' means that the corresponding partial derivative is non-negative, and '−' means that the corresponding partial derivative is non-positive. By definition, this is a consistent flow. This result could be obtained by combining the conclusions of Corollary 7 for each of the half-planes $x_2 \geq 0$ and $x_2 < 0$.

Theorem 11 *For some $t_1 \in I(\tilde{x})$, let $\tilde{\varphi}_{t_1}(\tilde{x}) = \tilde{x}_1$; then, $\forall\ t \geq t_1$ ($\forall\ t > 0$ if $t_1 = 0$), $t \in I(\tilde{x})$, $i \neq j$,*

$$\frac{\partial \varphi_t^j}{\partial x_i}(\tilde{x}) \begin{cases} > 0 \text{ if } (v_i, v_j) \text{ is positively consistently strongly connected} \\ \quad \text{in } G(\tilde{f}, \Gamma(\tilde{x})) \text{ and } G(\tilde{f}, \tilde{x}_1) \\ < 0 \text{ if } (v_i, v_j) \text{ is negatively consistently strongly connected} \\ \quad \text{in } G(\tilde{f}, \Gamma(\tilde{x})) \text{ and } G(\tilde{f}, \tilde{x}_1) \end{cases}$$

Proof: The proof begins in a similar way to the proof of Theorem 8. Suppose $(v_1, v_j), j \neq 1$, is consistently strongly connected in $G(\tilde{f}, \Gamma(\tilde{x}))$ and in $G(\tilde{f}, \tilde{x}_1)$, with $t_1 > 0$. Define the disjoint sets

$$\begin{aligned}
\mathcal{Q}_1 &= \{v_k : (v_1, v_k) \text{ is positively consistently strongly connected in } G(\tilde{f}, \Gamma(\tilde{x})) \\
&\quad \text{and in } G(\tilde{f}, \tilde{x}_1)\}, \\
\mathcal{R}_1 &= \{v_k : (v_1, v_k) \text{ is positively consistently strongly connected in } G(\tilde{f}, \Gamma(\tilde{x})) \\
&\quad \text{and not strongly connected in } G(\tilde{f}, \tilde{x}_1)\}, \\
\mathcal{Q}_2 &= \{v_k : (v_1, v_k) \text{ is negatively consistently strongly connected in } G(\tilde{f}, \Gamma(\tilde{x})) \\
&\quad \text{and in } G(\tilde{f}, \tilde{x}_1)\}, \\
\mathcal{R}_2 &= \{v_k : (v_1, v_k) \text{ is negatively consistently strongly connected in } G(\tilde{f}, \Gamma(\tilde{x})) \\
&\quad \text{and not strongly connected in } G(\tilde{f}, \tilde{x}_1)\}, \text{ and} \\
\mathcal{S} &= \{v_k : (v_1, v_k) \text{ is inconsistently strongly connected in } G(\tilde{f}, \Gamma(\tilde{x})) \\
&\quad \text{or } (v_1, v_k) \text{ is not strongly connected in } G(\tilde{f}, \Gamma(\tilde{x}))\}.
\end{aligned}$$

We relabel the vertices so that $v_1, \ldots, v_{q_1} \in \mathcal{Q}_1$, $v_{q_1+1}, \ldots, v_{r_1} \in \mathcal{R}_1$, $v_{r_1+1}, \ldots, v_{q_2} \in \mathcal{Q}_2$, $v_{q_2+1}, \ldots, v_{r_2} \in \mathcal{R}_2$, and $v_{r_2+1}, \ldots, v_n \in \mathcal{S}$. Performing the corresponding relabelling of $x_i, 2 \leq i \leq n$, puts the system (2.7) in the form

$$\dot{x}_i = f_i(x_1, \ldots, x_{r_2}), 1 \leq i \leq r_2,$$

$$\dot{x}_i = f_i(x_1, \ldots, x_n), r_2 + 1 \leq i \leq n.$$

We will show that for $t \geq t_1$,

$$\frac{\partial \phi_t^j}{\partial x_1}(\tilde{x}) \begin{cases} > 0, & 1 < j \leq q_1 \\ < 0, & r_1 + 1 \leq j \leq q_2 \end{cases}$$

For $1 \leq i \leq r_2$, we proceed as in the proof of Theorem 8. We have

$$\frac{\partial \dot{\phi}_t^j}{\partial x_1}(\tilde{x}) = \sum_{k=1}^{r_2} f_{i,k} \frac{\partial \phi_t^k}{\partial x_1}(\tilde{x}). \tag{2.16}$$

We can write (2.16) as $\dot{\tilde{Y}}(t) = M\tilde{Y}(t)$ where $\tilde{Y}$ is an r_2 vector and M is an $r_2 \times r_2$ matrix with $Y_i = \frac{\partial \phi_t^i}{\partial x_1}(\tilde{x})$ and $M_{i,k} = f_{i,k}$. Choosing $P = diag[P_i] = diag[1, \ldots, 1, -1, \ldots, -1]$, where P has r_1 entries of 1 and $r_2 - r_1$ entries of -1, means that PMP has non-negative off-diagonal entries.

Let $\tilde{Z} = P\tilde{Y}$. By Theorem 9, we know that $Z_k \geq 0$, $\forall\, t \in I(\tilde{x})$, $k \neq 1$, and that $Z_1 > 0$, $\forall\, t \in I(\tilde{x})$. We will prove that $Z_j(t) > 0$, $\forall\, t \geq t_1$, $1 < j \leq q_1$ and $r_1 + 1 \leq j \leq q_2$. Then the conclusion of the theorem would follow. As in the proof of Theorem 9, we have

$$\dot{\tilde{Z}} = (PMP)\tilde{Z} \Rightarrow \dot{\tilde{Z}} + \lambda\tilde{Z} = (PMP + \lambda)\tilde{Z} = N\tilde{Z},$$

where λ is chosen large enough so that $N = PMP + \lambda I \geq 0$. Solving for $\tilde{Z}$ in terms of the right hand side, we get

$$\tilde{Z}(t) = e^{\lambda t}\tilde{E} + \int_0^t e^{\lambda(s-t)} N\tilde{Z}(s)ds. \tag{2.17}$$

The proof will now proceed by induction on the length of the shortest directed path in $G(\tilde{f}, \tilde{x}_1)$ from v_1 to v_j. Suppose that a shortest directed path from v_1 to v_j in $G(\tilde{f}, \tilde{x}_1)$ has length 1. Then

$$\begin{aligned} Z_j(t) &= \int_0^t e^{\lambda(s-t)} \sum_{l=1}^{r_2} N_{jl} Z_l ds \\ &\geq \int_0^t e^{\lambda(s-t)} N_{j1} Z_1 ds. \end{aligned} \tag{2.18}$$

Since $N_{j1}(t_1) = P_j f_{j,1}(\tilde{x}_1)P_1 > 0$, we conclude that $Z_j(t) > 0$, for $t \geq t_1$.

Now suppose the result is true if the shortest directed path has length m. We consider the case when the shortest directed path from v_1 to v_j in $G(\tilde{f}, \tilde{x}_1)$ has length $m+1$. Suppose the intermediate vertices are $v_{k_1}, \ldots, v_{k_m}$, with v_{k_1} adjacent to v_1, v_{k_l} adjacent to $v_{k_{l+1}}$, $1 \leq l \leq m-1$, and v_{k_m} adjacent to v_{k_j}. Note that each k_l satisfies either $1 < k_l \leq q_1$ or $r_1 + 1 \leq k_l \leq q_2$. Then

$$\begin{aligned} Z_j(t) &= \int_0^t e^{\lambda(s-t)} \sum_{l=1}^{r_2} N_{jl} Z_l ds \\ &\geq \int_0^t e^{\lambda(s-t)} N_{jk_m} Z_{k_m} ds. \end{aligned} \tag{2.19}$$

Again, $N_{jk_m}(t_1) = P_j f_{j,k_m}(\tilde{x}_1)P_{k_m} > 0$. In order for a shortest directed path from v_1 to v_j to have length $m+1$, a shortest path from v_1 to v_{k_m} must have length m; hence, $Z_{k_m}(t) > 0$, $t \geq t_1$, by the inductive hypothesis. Thus, $Z_j(t) > 0$, $t \geq t_1$. The proof by induction is now complete.

The case $t_1 = 0$ is argued in exactly the same way. □

Theorem 11 requires some knowledge of the solution trajectory in order to be useful. The following corollary gives a result which does not require any information about the solution trajectory; it is most useful in applications.

Corollary 12 *If (v_i, v_j) is positively (negatively) consistently strongly connected in $G(\tilde{f}, \Omega)$, $\tilde{x}(0) \in \bar{\Omega}$, $\tilde{x}(t) \in \Omega$ for $t > 0$, then*

$$\frac{\partial \varphi_t^j}{\partial x_i}(\tilde{x}) > 0 \ (< 0) \text{ for } t > 0.$$

A particular, well-known case of Corollary 12 is stated by M. Hirsch in [11] and [13]:

Corollary 13 *If $\tilde{f}$ is a cooperative vector field and $D\tilde{f}(\tilde{x})$ is also irreducible for all $\tilde{x}$, then $\{\varphi_t\}$ has positive derivatives.*

Recall that a cooperative vector field is one that satisfies $f_{i,j}(\tilde{x}) \geq 0$, $\forall\ \tilde{x} \in \Omega$, $\forall\ i, j,\ i \neq j$. $D\tilde{f}(\tilde{x})$ is irreducible means that $G(\tilde{f}, \Omega)$ is strongly connected, that there is a directed path from any vertex to any other vertex.

It is perhaps important to make some comments on Corollary 13. The proofs in [11] and [13] are incorrect (verified by M. Hirsch). References to the (incorrect) proof pervade the literature, but a careful search reveals the following. [1] contains a result that can be used to establish Corollary 13. [24] contains a related result with stricter hypotheses. [34] refers to [13] without mentioning the incorrect proof. [33] and [35] contain a correct proof. [19] contains a proof of Theorem 11 from which the corollary immediately follows. In [41], a generalization of Corollary 13, with weaker hypotheses, is given.

2.2.4 An Algorithm For Complicated Multigraphs

It may be difficult to check the conditions of the previous theorems, so we present an algorithm for dealing with complicated multigraphs. We will require the following graph theory results.

Lemma 14 *Delete the sources and sinks (and their incident edges) from $G(\tilde{f}, \Omega)$. Repeat this procedure on the resulting graph as long as there are sources or sinks remaining. The result is the* empty graph *(no vertices or edges) if and only if the original graph has no directed cycles. There is a directed cycle involving some of the vertices that survive this procedure.*

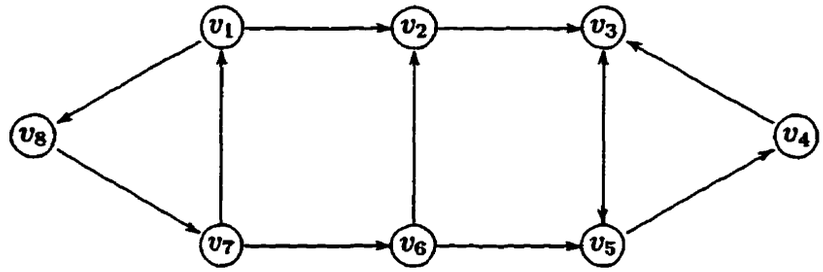


Figure 4: Example graph for Lemma 14.

Proof: See [27], pages 18 and 332. □

Remark: Not all vertices that survive the process outlined in Lemma 14 are necessarily involved in a directed cycle. Consider the graph in Figure 4. The graph contains no sources or sinks, yet neither vertex v_2 nor vertex v_6 is involved in a directed cycle.

Lemma 15 *Let G be a signed, directed, connected graph all of whose vertices are contained in the same directed cycle $\mathcal{C}$. Delete one edge from $\mathcal{C}$ to obtain a directed path $\mathcal{P}$. In turn, add the edges in G which are not in $\mathcal{P}$, producing (ignoring leftover edges) either a directed cycle or two co-terminal directed paths, where the second possibility will be called a directed bicycle. Let $\mathcal{S}$ denote the set of these directed cycles and directed bicycles. Then the vertices in G are consistently strongly connected if and only if all of the directed cycles and directed bicycles in $\mathcal{S}$ are positive.*

Proof: If all of the directed cycles and directed bicycles in $\mathcal{S}$ are positive, then all cycles in a fundamental set of cycles of the associated undirected graph G_u are positive, implying that the vertices are consistently strongly connected. If a directed cycle or a directed bicycle is negative, then (at least) two of the vertices of G are inconsistently strongly connected. Since the vertices are strongly connected by the di-

rected cycle C, it would then follow that any two vertices are inconsistently strongly connected. □

A directed graph can be represented by an adjacency matrix. For such a graph with n vertices, the adjacency matrix $\mathcal{A}$ is an $n \times n$ matrix with ij^{th} element a_{ij} where $a_{ij} = 1$ if there is a directed edge from vertex v_i to vertex v_j and $a_{ij} = 0$ otherwise. Since there are no loops (edges with the same start and end vertices), $a_{ii} = 0$.

The powers of $\mathcal{A}$ give information about the walks in the graph. There are $(\mathcal{A}^p)_{ij}$ directed walks of length p from v_i to v_j.

We will need a similar matrix representation for a *weighted, directed graph*, a directed graph in which each edge has a value (not necessarily numeric) associated with it. With this in mind, we will represent a weighted, directed graph by a matrix $\mathcal{W}$ which captures all of the connections and weights. For such a graph with n vertices, the $n \times n$ matrix $\mathcal{W}$ has ij^{th} element w_{ij} equal to the weight of the directed edge from v_i to v_j. If there is no directed edge from v_i to v_j, we set $w_{ij} = 0$. In a graph with no loops, $w_{ii} = 0$.

In general, $G(\tilde{f}, \Omega)$ can be a complex graph. This algorithm offers a way to collapse $G(\tilde{f}, \Omega)$ to a weighted, directed graph, $G_w(\tilde{f}, \Omega)$, that can be analyzed systematically; the analysis can then be extended back to $G(\tilde{f}, \Omega)$.

Step I: Collapse the signed, directed multigraph $G(\tilde{f}, \Omega)$ to a signed, directed multigraph with no directed cycles, $G_\emptyset(\tilde{f}, \Omega)$. We perform the following iterative procedure, eliminating one directed cycle with each iteration.

(i) Let the iteration counter be k. Let $k = 0$ to start and let $G_0(\tilde{f}, \Omega) = G(\tilde{f}, \Omega)$. Each iteration produces $G_{k+1}(\tilde{f}, \Omega)$ from $G_k(\tilde{f}, \Omega)$.

(ii) Let $\mathcal{V}_k = \{v_1, \ldots, v_{n_k}, v_1^{\circ}, \ldots, v_{m_k}^{\circ}\}$ represent the vertex set of $G_k(\tilde{f}, \Omega)$ and let $m_k = 0$ to start. The notation will become clearer shortly.

(iii) Find a directed cycle in $G_k(\tilde{f}, \Omega)$, by using Lemma 14 if necessary. Let $\mathcal{V}_k^{\circ}$ be the set of vertices of $G_k(\tilde{f}, \Omega)$ in the directed cycle. If no directed cycle exists in $G_k(\tilde{f}, \Omega)$, then Step I is completed.

(iv) $G_{k+1}(\tilde{f}, \Omega)$ has vertex set $\mathcal{V}_{k+1} = \mathcal{V}_k \setminus \mathcal{V}_k^{\circ} + \{v_k^{\circ}\}$. That is, the vertices of $G_k(\tilde{f}, \Omega)$ not in the directed cycle remain vertices of $G_{k+1}(\tilde{f}, \Omega)$, but vertices of $G_k(\tilde{f}, \Omega)$ in the directed cycle are collapsed into one vertex, v_k°, in $G_{k+1}(\tilde{f}, \Omega)$.

We will say that vertices of $\mathcal{V}_k^{\circ}$ are contained in v_k°. For $v_i \in G(\tilde{f}, \Omega)$, if $v_i \in \mathcal{V}_k^{\circ}$ in $G_k(\tilde{f}, \Omega)$ then we will write $v_i \in v_k^{\circ}$ in $G_{k+1}(\tilde{f}, \Omega)$. If $v_i \in v_{k_1}^{\circ} \in v_{k_2}^{\circ} \in \ldots \in v_{k_l}^{\circ}$ in $G_k(\tilde{f}, \Omega)$ then we will write $v_i \in v_{k_l}^{\circ}$ in $G_{k+1}(\tilde{f}, \Omega)$.

It will be important to keep track of which vertices are collapsed and which vertices they (ultimately) contain.

If the vertices in $\mathcal{V}_k^{\circ}$ are connected by other edges, Lemma 15 must be applied to determine whether the vertices are consistently strongly connected. This is only necessary if the directed cycle is positive and contains no negative vertices. v_k° is labelled positive (negative) if the vertices are consistently (inconsistently) strongly connected. If the directed cycle contains a negative vertex then v_k° is negative.

(v) To avoid confusion when signing the edges of $G_{k+1}(\tilde{f}, \Omega)$, we pick a vertex of the directed cycle, say $v_k^* \in \mathcal{V}_k^{\circ}$; it is sensible

to pick the vertex of highest degree. Now, suppose there is a directed edge from v_i to v_j in $G_k(\tilde{f}, \Omega)$ (v_i and v_j might be collapsed vertices).

(a) if $v_i, v_j \in \mathcal{V}_k \setminus \mathcal{V}_k^{\circ}$ then a directed edge of the same sign from v_i to v_j is drawn in $G_{k+1}(\tilde{f}, \Omega)$.

(b) if $v_i \in V_k \setminus \mathcal{V}_k^{\circ}$ and $v_j \in \mathcal{V}_k^{\circ}$ then a directed edge from v_i to v_k° is drawn in $G_{k+1}(\tilde{f}, \Omega)$. This edge in $G_{k+1}(\tilde{f}, \Omega)$ has the same sign as the shortest path from v_i to v_k^* in $G_k(\tilde{f}, \Omega)$.

(c) if $v_i \in \mathcal{V}_k^{\circ}$ and $v_j \in \mathcal{V}_k \setminus \mathcal{V}_k^{\circ}$ then a directed edge from v_k° to v_j is drawn in $G_{k+1}(\tilde{f}, \Omega)$. This edge in $G_{k+1}(\tilde{f}, \Omega)$ has the same sign as the shortest path from v_k^* to v_j in $G_k(\tilde{f}, \Omega)$.

(d) if $v_i, v_j \in \mathcal{V}_k^{\circ}$ then no edge is drawn in $G_{k+1}(\tilde{f}, \Omega)$.

Edges with the same sign and direction need not be drawn more than once. It is interesting to note that if the directed cycle is negative, one need not be careful when signing the edge in (b) or (c) above since the next step of the algorithm will erase the signs of these edges.

(vi) Increase k by 1 and return to (ii). The process is completed when no directed cycle is found in $G_k(\tilde{f}, \Omega)$ in (iii).

The result of this process is in general a signed, directed, tripartite (three different kinds of vertices) multigraph with no directed cycles, $G_{\emptyset}(\tilde{f}, \Omega)$, but the only possible parallel edges will be edges of opposite sign in the same direction.

Step II: Transform the signed, directed, tripartite multigraph produced in Step I, $G_{\emptyset}(\tilde{f},\Omega)$, into a weighted, directed, tripartite graph, $G_w(\tilde{f},\Omega)$, on the same vertices, using the following steps in order:

(i) Replace parallel or anti-parallel edges of opposite sign by a single edge in the same direction weighted *.

(ii) Assign a weight of * to any edge adjacent to a negative vertex. This includes both incoming and outgoing edges.

(iii) Assign a weight of +1 (−1) to any positive (negative) edge.

The * weighting is given to edges which carry an inconsistency. A negative vertex involves vertices of $G(\tilde{f},\Omega)$ that are connected by walks of both signs; hence, edges adjacent to a negative vertex in the collapsed graph carry this inconsistency in whichever direction they point.

Step III: Making no distinction between collapsed vertices and original vertices, construct the matrix $\mathcal{W}$ associated with $G_w(\tilde{f},\Omega)$. $\mathcal{W}$ has possible entries of 0, +1, -1, and *. Let the vertices of $G_w(\tilde{f},\Omega)$ be $\{\bar{v}_1,\ldots,\bar{v}_{n_w}\}$.

Step IV: We define

$$\begin{aligned} |*| &= *, \\ a+* &= *,\ a=0,-1,1,*, \\ a\cdot* &= *,\ a=-1,1,*,\ \text{and} \\ 0\cdot* &= 0. \end{aligned}$$

Since $G_w(\tilde{f},\Omega)$ has no directed cycles, any directed walk can have length at most equal to n_w-1; hence, $\mathcal{W}^{n_w}$ is the zero matrix.

We compare entries in $|W^p|$ and $|W|^p$ for $2 \leq p \leq n_w - 1$, where $|W|_{ij} = |W_{ij}|$. For any $i \neq j$, there are three possibilities:

(i) If $|W|^p_{ij} = 0$ for each p then $(\bar{v}_i, \bar{v}_j)$ is not strongly connected in $G_w(\tilde{f}, \Omega)$.

(ii) Otherwise, if $|W|^p_{ij} = |W^p|_{ij} \neq *$ and $sign(W^p)_{ij}$ is the same for each p then $(\bar{v}_i, \bar{v}_j)$ is consistently strongly connected in $G_w(\tilde{f}, \Omega)$; the sign of the connection can be determined by finding a directed path in the graph from v_i to v_j or by examining the non-zero sign of W^p_{ij} for some p, $2 \leq p \leq n_w - 1$.

(iii) All other vertex pairs $(\bar{v}_i, \bar{v}_j)$ are inconsistently strongly connected in $G_w(\tilde{f}, \Omega)$.

Remark: We are extending our earlier definitions of consistently and inconsistently strongly connected for $G_k(\tilde{f}, \Omega)$ and $G_w(\tilde{f}, \Omega)$. A vertex pair in $G_k(\tilde{f}, \Omega)$ is consistently (inconsistently) strongly connected if all directed walks are of the same sign and none includes a negative vertex (if there are directed walks of each sign or a directed walk includes a negative vertex). A vertex pair in $G_w(\tilde{f}, \Omega)$ is consistently (inconsistently) strongly connected if all directed walks are of the same sign and none includes * weighted edges (if there are directed walks of each sign or a directed walk includes a * weighted edge).

Before proceeding, we introduce the following lemma.

Lemma 16 *Let $v_i \in v_k^{\circlearrowright}$ in $G_{\emptyset}(\tilde{f}, \Omega)$. Then $v_k^{\circlearrowright}$ contains only the vertices v_j of $G(\tilde{f}, \Omega)$ where there is both a directed walk from v_i to v_j and one from v_j to v_i in $G(\tilde{f}, \Omega)$.*

Proof: We proceed by induction on the counter k in Step I of the algorithm. Suppose $k = 1$; then, the vertices contained in v_k° are vertices in $G(\tilde{f}, \Omega)$ that comprise a directed cycle in $G(\tilde{f}, \Omega)$ and we are done.

Suppose the claim holds for $k = r - 1$. Consider the case of $k = r$. The claim holds for $G_{r-1}(\tilde{f}, \Omega)$; hence, in $G_r(\tilde{f}, \Omega)$ we need only consider the newly added collapsed vertex, v_r°, which corresponds to a directed cycle in $G_{r-1}(\tilde{f}, \Omega)$. There are several possibilities for two distinct vertices $v_i, v_j \in v_r^{\circ}$:

(i) v_i and v_j are vertices of $G_{r-1}(\tilde{f}, \Omega)$; then the claim holds by assumption.

(ii) v_i and v_j are contained in the same collapsed vertex of $G_{r-1}(\tilde{f}, \Omega)$; then the claim holds by assumption.

(iii) v_i and v_j are contained in the different collapsed vertices of $G_{r-1}(\tilde{f}, \Omega)$, say v_i° and v_j° respectively; then there is a directed walk from v_i (v_j) to each other vertex in v_i° (v_j°) and to v_i (v_j) from each other vertex in v_i° (v_j°). Using those directed walks and the connections between v_i° and v_j° in $G_{r-1}(\tilde{f}, \Omega)$, one can construct directed walks in each direction between v_i and v_j.

(iv) only one of v_i or v_j is contained in a collapsed vertex of $G_{r-1}(\tilde{f}, \Omega)$. Similar reasoning to the above works in this case.

□

Lemma 16 allows us to state that $G_{\emptyset}(\tilde{f}, \Omega)$ and hence $G_w(\tilde{f}, \Omega)$ are unique for a given $G(\tilde{f}, \Omega)$, regardless of the order in which directed cycles are identified in Step I(iii). Furthermore, we will see that all vertices of $G(\tilde{f}, \Omega)$ in a positive (negative) collapsed vertex of $G_w(\tilde{f}, \Omega)$ are consistently (inconsistently) strongly

connected. These results, drawing conclusions on the connections in $G(\tilde{f},\Omega)$ from the connections in $G_w(\tilde{f},\Omega)$, will be presented in the next theorem.

For convenience, we first introduce the following notation. Let the vertices of $G(\tilde{f},\Omega)$ be $\{v_1,\ldots,v_n\}$. If (v_i,v_j) is not strongly connected (positively consistently strongly connected, negatively consistently strongly connected, inconsistently strongly connected) in $G(\tilde{f},\Omega)$, then we write $(v_i,v_j)\in \mathcal{S}_1(\mathcal{S}_2,\mathcal{S}_3,\mathcal{S}_4)$. We will use the same notation for vertex pairs, $(\bar{v}_i,\bar{v}_j)$, of $G_w(\tilde{f},\Omega)$.

The set of positive (negative) collapsed vertices in $G_w(\tilde{f},\Omega)$ will be denoted by $\mathcal{V}^+$ ($\mathcal{V}^-$). We let $\mathcal{V}^{\mathcal{O}}=\mathcal{V}^+\cup\mathcal{V}^-$ and let $\mathcal{V}^0$ denote the set of vertices in $G_w(\tilde{f},\Omega)$ that are also vertices in $G(\tilde{f},\Omega)$. The next theorem follows from the work of this section:

Theorem 17 *After performing the algorithm in Steps I–IV, all vertex connections in $G_w(\tilde{f},\Omega)$ can be classified according to the following rules:*

1. *Let $v_i,v_j\in\bar{v}_l\in\mathcal{V}^{\mathcal{O}}$. If $\bar{v}_l\in\mathcal{V}^+(\bar{v}_l\in\mathcal{V}^-)$, then $(v_i,v_j)\in\mathcal{S}_2\cup\mathcal{S}_3$ ($\mathcal{S}_4$). In the first case, examining the sign of a directed path from v_i to v_j in $G(\tilde{f},\Omega)$ determines whether $(v_i,v_j)\in\mathcal{S}_2\cup\mathcal{S}_3$.*

2. *Let $v_i\in\bar{v}_l$, $v_j\in\bar{v}_m$, $\bar{v}_l\neq\bar{v}_m$, $\bar{v}_l,\bar{v}_m\in\mathcal{V}^{\mathcal{O}}$. If $(\bar{v}_l,\bar{v}_m)\in\mathcal{S}_1$ ($\mathcal{S}_4$) in $G_w(\tilde{f},\Omega)$, then $(v_i,v_j)\in\mathcal{S}_1$ ($\mathcal{S}_4$) in $G(\tilde{f},\Omega)$. If $(\bar{v}_l,\bar{v}_m)\in\mathcal{S}_2\cup\mathcal{S}_3$ in $G_w(\tilde{f},\Omega)$, then $(v_i,v_j)\in\mathcal{S}_2\cup\mathcal{S}_3$ in $G(\tilde{f},\Omega)$. If $\bar{v}_l\in\mathcal{V}^-$ or $\bar{v}_m\in\mathcal{V}^-$ then only $(v_i,v_j)\in\mathcal{S}_1$ or $(v_i,v_j)\in\mathcal{S}_4$ are possible.*

3. *Let $v_i=\bar{v}_l\in\mathcal{V}^0$, $v_j\in\bar{v}_m\in\mathcal{V}^{\mathcal{O}}$. If $(\bar{v}_l,\bar{v}_m)\in\mathcal{S}_1$ ($\mathcal{S}_4$) in $G_w(\tilde{f},\Omega)$, then $(v_i,v_j)\in\mathcal{S}_1$ ($\mathcal{S}_4$) in $G(\tilde{f},\Omega)$. If $(\bar{v}_l,\bar{v}_m)\in\mathcal{S}_2\cup\mathcal{S}_3$ in $G_w(\tilde{f},\Omega)$, then $(v_i,v_j)\in\mathcal{S}_2\cup\mathcal{S}_3$ in $G(\tilde{f},\Omega)$. If $\bar{v}_m\in\mathcal{V}^-$ then only $(v_i,v_j)\in\mathcal{S}_1$ or $(v_i,v_j)\in\mathcal{S}_4$ are possible.*

4. Let $v_i = \bar{v}_l \in \mathcal{V}^0$, $v_j = \bar{v}_m \in \mathcal{V}^0$. *If* $(\bar{v}_l, \bar{v}_m) \in \mathcal{S}_1$ $(\mathcal{S}_4)$ *in* $G_w(\tilde{f}, \Omega)$, *then* $(v_i, v_j) \in \mathcal{S}_1$ $(\mathcal{S}_4)$ *in* $G(\tilde{f}, \Omega)$. *If* $(\bar{v}_l, \bar{v}_m) \in \mathcal{S}_2 \cup \mathcal{S}_3$ *in* $G_w(\tilde{f}, \Omega)$, *then* $(v_i, v_j) \in \mathcal{S}_2 \cup \mathcal{S}_3$ *in* $G(\tilde{f}, \Omega)$.

Proof: We first observe that at each iteration of Step 1(v) the consistency or inconsistency of pairs of vertices are maintained in the following sense. In case (a), (v_i, v_j) is consistently (inconsistently) strongly connected in $G_k(\tilde{f}, \Omega)$ if and only if (v_i, v_j) is consistently (inconsistently) strongly connected in $G_{k+1}(\tilde{f}, \Omega)$. In case (b), (v_i, v_j) is consistently (inconsistently) strongly connected in $G_k(\tilde{f}, \Omega)$ if and only if $(v_i, v_k^{\diamond})$ is consistently strongly connected in $G_{k+1}(\tilde{f}, \Omega)$ and $v_k^{\diamond}$ is positive (inconsistently strongly connected, i.e. either $v_k^{\diamond}$ is negative or there are walks of each sign from v_i to $v_k^{\diamond}$). In case (c), (v_i, v_j) is consistently (inconsistently) strongly connected in $G_k(\tilde{f}, \Omega)$ if and only if $(v_k^{\diamond}, v_j)$ is consistently strongly connected in $G_{k+1}(\tilde{f}, \Omega)$ and $v_k^{\diamond}$ is positive (inconsistently strongly connected, i.e. either $v_k^{\diamond}$ is negative or there are walks of each sign from $v_k^{\diamond}$ to v_j). In case (d), (v_i, v_j) is consistently (inconsistently) strongly connected in $G_k(\tilde{f}, \Omega)$ if and only if $v_k^{\diamond}$ is positive (negative).

Case 1. We now see that only consistently strongly connected vertices of $G(\tilde{f}, \Omega)$ will be contained in a positive vertex in $G_w(\tilde{f}, \Omega)$ and that all vertices contained in a negative vertex in $G_w(\tilde{f}, \Omega)$ are inconsistently strongly connected. The result follows.

Cases 2, 3, and 4. In each case, $(\bar{v}_l, \bar{v}_m) \in \mathcal{S}_1$ $(\mathcal{S}_4) \Rightarrow (v_i, v_j) \in \mathcal{S}_1$ $(\mathcal{S}_4)$ follows immediately. $(\bar{v}_l, \bar{v}_m) \in \mathcal{S}_2 \cup \mathcal{S}_3 \Rightarrow (v_i, v_j) \in \mathcal{S}_2 \cup \mathcal{S}_3$ follows as well, with the sign of the connection being determined by any path from v_i to v_j in $G(\tilde{f}, \Omega)$. This information may be buried in the collapsed vertices of $G_w(\tilde{f}, \Omega)$.

□

Finally, we can connect Theorem 17 with Theorem 9 of the previous section to draw conclusions on the signs of partial derivatives.

We should note that two inconsistently strongly connected vertices could still correspond to solution components which have some monotonicity with respect to each other. The theory of this chapter does not help us; but the next chapter makes some inroads (see example 12).

Before presenting several examples, we offer a method for determining when the positivity assumption is satisfied for a class of chemical kinetics problems; this is highlighted in Example 10.

2.3 Positivity For a Class Of Reaction Mechanisms

In general, a careful analysis is required when seeking initial sets that satisfy the positivity assumption (see Section 1.1). A graph theoretic approach is given in [39]; it determines which concentrations will be positive for $t > 0$ when a particular set of species is present initially. For a certain class of reaction mechanisms, the analysis simplifies and we present it here in the following theorem which determines the smallest possible sets of species which must be present initially. Note that this theorem uses a different graph than the one in [39].

Theorem 18 Consider a reaction mechanism which involves reactions of only two types:

(i) the product species are a subset of the reactant species, or

(ii) there is a single reactant species.

Suppose that the reactions of type (ii) involve species $A_1, \ldots, A_m$. We draw a directed multigraph G_+ with m vertices, $v_1, \ldots, v_m$, with vertex v_i corresponding to species A_i. Directed edges are drawn from vertex v_i to vertex v_j, $i \neq j$, if a reaction of type (ii) with species A_i as a reactant produces species A_j. We call each source vertex of G_+ (a vertex with no incoming edges) and each strongly connected subgraph of G_+ with no incoming edges an initial group.

In each initial group, at least one species must be present initially to guarantee positivity of all species for $t > 0$. In addition, if there is a species in a reaction of type (i) that does not occur in any of the reactions of type (ii) (as a reactant or as a product), it must also be present initially.

Proof: [39] provides a labelling scheme to determine which species will be present for $t > 0$ given that certain species are present initially. The species which are present initially are labelled with a 0. In the first step of the labelling process, the products of reactions with only 0-labelled species as reactants are labelled 1. The process continues in this manner, with unlabelled products of reactions with labelled reactants getting the label for the current step. At the end, all labelled species will be present for $t > 0$.

In our set-up, each vertex of G_+ is either in an initial group or connected to an initial group by a directed walk from the group to the vertex. The vertices of G_+ correspond to all of the species in the chemical system except for species in a reaction of type (i) that do not occur in any of the reactions of type (ii).

If no species in an initial group is present initially, it is clear that none of the species in that group will ever be present. If a species in a reaction of type (i) that

does not occur in any of the reactions of type (ii) is not present initially, then it is clear that this species will never be present.

If at least one species in an initial group is present initially, then, by the strong connectedness of the group in G_+, all of the species will be present for $t > 0$. Any vertices that are connected to an initial group by a directed walk (from the group to the vertex) will also be present for $t > 0$. If this is the case for each initial group, then we need only additionally insure that any species in a reaction of type (i) that do not occur in any of the reactions of type (ii) are also initially present. In this case, all species are present for $t > 0$. □

2.4 Examples

Example 4 (Neural Networks): A standard equation to model a neural network is

$$\dot{x}_i = H_i(x_i, s_i) = F_i(x_1, \ldots, x_n),\ i = 1, \ldots, n, \tag{2.20}$$

where

$$s_i = \sum_j W_{ij} g_j(x_j),\ W_{ij} \text{ constant },\ \frac{\partial H_i}{\partial s_i} \geq 0,\ \text{and}\ \frac{dg_j}{dx_j} \geq 0.$$

See [16] or [14] for more details. Two mathematically interesting cases are (i) *excitatory nets*, and (ii) *even-loop nets*. In case (i), $W_{ij} \geq 0$, $i \neq j$, so $G(\tilde{F}, \Omega)$ has only positive edges; by Theorem 5, the system induces an order preserving flow (monotone flow, in this case). In case (ii), every directed cycle in $G(\tilde{F}, \Omega)$ is positive, so Theorem 5 applies again. In both cases, one can determine whether partial derivatives are negative, positive, or zero for all time by examining $G(\tilde{F}, \Omega)$.

Example 5 (Chemical Kinetics): Consider a chemical reaction mechanism con-

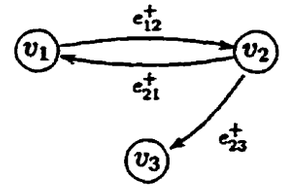


Figure 5: $G(\tilde{f}, \Omega)$ for Example 6

sisting of reactions of the form

$$\alpha_{il} A_i \xrightarrow{k_l} \beta_{jl} A_j.$$

It is easily seen that the corresponding signed, directed multigraph has only positive edges; hence, by Theorem 5, the system induces an order preserving flow (monotone flow, in this case). By investigating the directed graph $G(\tilde{f}, \Omega)$, we can determine if each partial derivative is positive or zero for all positive time.

Example 6 (Chemical Kinetics): In [30], the reaction mechanism

$$2A_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} A_2 \xrightarrow{k_2} A_3$$

was considered and the signs of partial derivatives of concentrations with respect to initial concentrations were given without proof. This mechanism induces the system of differential equations,

$$\begin{aligned} \dot{x}_1 &= -k_1 x_1^2 + k_{-1} x_2, \\ \dot{x}_2 &= -(k_2 + k_{-1}) x_2 + k_1 x_1^2, \text{ and} \\ \dot{x}_3 &= k_2 x_2. \end{aligned}$$

Figure 5 presents the multigraph $G(\tilde{f}, \Omega)$ for this mechanism. The positivity as-

sumption will hold if $\tilde{x}_0 \in \mathcal{S}_0 = \{x_1 + x_2 > 0\}$. In such a situation, Table 1 summarizes the signs of the partial derivatives, where '++' means positive for $t \geq 0$, '+' means positive for $t > 0$, and '0' means the derivative is zero for all time.

	$x_1(t)$	$x_2(t)$	$x_3(t)$
$x_{1,0}$	++	+	+
$x_{2,0}$	+	++	+
$x_{3,0}$	0	0	++

Table 1: Signs of concentrations with respect to changes in initial concentrations for Example 6.

Example 7 (Chemical Kinetics): Consider the simple bimolecular reaction

$$A_1 + A_2 \xrightarrow{k} A_3,$$

which leads to the system of differential equations

$$\begin{aligned} \dot{x}_1 &= -kx_1x_2, \\ \dot{x}_2 &= -kx_1x_2, \text{ and} \\ \dot{x}_3 &= +kx_1x_2. \end{aligned}$$

We construct the multigraph in Figure 6, where, as usual, dashed edges have negative sign and solid edges have positive sign. It is easily seen that the graph $G_u(\tilde{f}, \Omega)$ consists of a single negative cycle, confirming by Theorem 5 that the reaction does not induce an order preserving flow. Using Theorem 9, we can still conclude that if the positivity assumption is satisfied, Table 2 gives the signs of the partial derivatives of concentrations with respect to initial concentrations, where the table entries have the same meaning as in Example 6, with '−' meaning that the derivative is

Figure 6: $G(\tilde{f}, \Omega)$ and $G_u(\tilde{f}, \Omega)$ for Example 7

negative for $t > 0$. Blank entries mean that we have not determined the sign of the derivative. The positivity assumption will be satisfied if $\tilde{x}_0 \in \mathcal{S}_0 = \{x_1 > 0, x_2 > 0\}$.

	$x_1(t)$	$x_2(t)$	$x_3(t)$
$x_{1,0}$	$++$	$-$	
$x_{2,0}$	$-$	$++$	
$x_{3,0}$	0	0	$++$

Table 2: Signs of concentrations with respect to changes in initial concentrations for Example 7.

Example 8 (Chemical Kinetics): Consider the same bimolecular reaction of Example 7, with the reaction now being reversible:

$$A_1 + A_2 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} A_3,$$

This mechanism leads to the system of differential equations

$$\begin{aligned} \dot{x}_1 &= -k_1 x_1 x_2 + k_{-1} x_3, \\ \dot{x}_2 &= -k_1 x_1 x_2 + k_{-1} x_3, \text{ and} \\ \dot{x}_3 &= +k_1 x_1 x_2 - k_{-1} x_3, \end{aligned}$$

and the corresponding multigraph is given in Figure 7, where dashed edges have negative sign and solid edges have positive sign. As in Example 7, $G_u(\tilde{f}, \Omega)$ consists

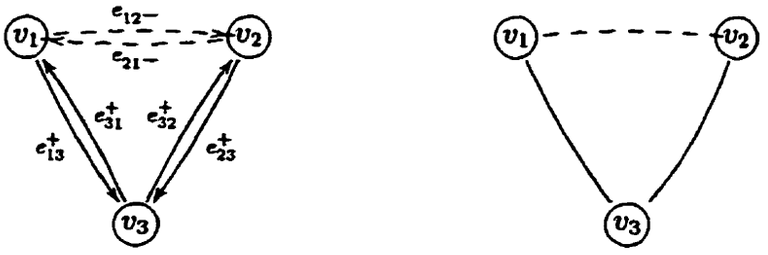


Figure 7: $G(\tilde{f}, \Omega)$ and $G_u(\tilde{f}, \Omega)$ for Example 8

of a single negative cycle, so the reaction does not induce an order preserving flow. This time, however, we cannot draw any conclusions on the partial derivatives since every pair of vertices is inconsistently strongly connected in $G(\tilde{f}, \Omega)$. We will return to this example in the next chapter.

Example 9 (Epidemiology): Consider the following mechanism for the SIS epidemic model (see [7]):

$$S + I \xrightarrow{\beta} 2I$$

$$I \xrightarrow{\gamma} S$$

This mechanism describes an epidemic in which susceptibles (S) meet infectives (I) to produce two infectives at a positive rate of β, while infectives recover without immunity at a positive rate of γ.

Let the time dependent populations of susceptibles and infectives be denoted by the variables $x_1(t)$ and $x_2(t)$, respectively. The law of mass-action gives the system of differential equations

$$\dot{x}_1(t) = f_1(x_1, x_2) = \gamma x_2(t) - \beta x_1(t) x_2(t), \text{ and} \tag{2.21}$$

$$\dot{x}_2(t) = f_2(x_1, x_2) = -\gamma x_2(t) + \beta x_1(t) x_2(t). \tag{2.22}$$

To insure that we have a non-trivial situation, we require that $x_2(0) > 0$.

This problem was analyzed in [18] and [31]; problem-specific arguments were needed to obtain monotonicity results for this model. The results are summarized in the Table 3, which gives the sign of a partial derivative of a population with respect to an initial population. The */+ entry in Table 3 means that the partial derivative is always positive for the case $x_1(0) \leq \frac{\gamma}{\beta}$, and of both signs for the case $x_1(0) > \frac{\gamma}{\beta}$. The $-/+$ entry means that the partial derivative is always positive for the case $x_1(0) < \frac{\gamma}{\beta}$, always negative for the case $x_1(0) > \frac{\gamma}{\beta}$, and identically zero for all time in the case $x_1(0) = \frac{\gamma}{\beta}$. Before attempting to obtain the monotonicity

	$x_1(t)$	$x_2(t)$
$x_1(0)$	*/+	+
$x_2(0)$	$-/+$	+

Table 3: Behaviour of populations with respect to changes in initial populations for the SIS epidemic model for Example 9.

results of Table 3 using the methods of Chapter 2, we first notice that

(i) if $x_1(0) < \frac{\gamma}{\beta}$ then $x_1(t) < \frac{\gamma}{\beta}$, $\forall\, t \geq 0$,

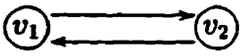
(ii) if $x_1(0) = \frac{\gamma}{\beta}$ then $x_1(t) = \frac{\gamma}{\beta}$, $\forall\, t \geq 0$, and

(iii) if $x_1(0) > \frac{\gamma}{\beta}$ then $x_1(t) > \frac{\gamma}{\beta}$, $\forall\, t \geq 0$.

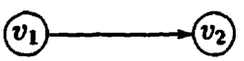
The Jacobian matrix for this problem is

$$D\tilde{f} = \begin{pmatrix} -\beta x_2 & \gamma - \beta x_1 \\ \beta x_2 & -\gamma + \beta x_1 \end{pmatrix}.$$

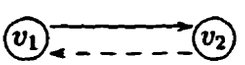
For case (i), applying the theory of this chapter, we draw the directed graph



which consists of a single positive directed cycle. For case (ii), applying the theory of this chapter, we draw the graph



which also contains no inconsistencies. These two graphs give us the '+' entries in Table 3 in the case $x_1(0) \leq \frac{\gamma}{\beta}$. For case (iii), applying the theory of this chapter, we draw the graph



which consists of a single negative directed cycle. The theory of this chapter takes us no further.

Example 10 (Chemical Kinetics): The Michaelis-Menten reactions of enzyme kinetics can be written

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\rightarrow} P + E$$

where E, S, ES, and P are the enzyme, substrate, complex, and product, respectively. We will denote the concentrations of E, S, ES, and P by $x_1(t)$, $x_2(t)$, $x_3(t)$, and $x_4(t)$, respectively; mass action chemical kinetics yields the system

$$\dot{x}_1(t) = f_1(x_1, x_2, x_3, x_4) = -k_1 x_1 x_2 + (k_{-1} + k_2) x_3, \tag{2.23}$$

$$\dot{x}_2(t) = f_2(x_1, x_2, x_3, x_4) = -k_1 x_1 x_2 + k_{-1} x_3, \tag{2.24}$$

$$\dot{x}_3(t) = f_3(x_1, x_2, x_3, x_4) = k_1 x_1 x_2 - (k_{-1} + k_2) x_3, \text{ and} \tag{2.25}$$

$$\dot{x}_4(t) = f_4(x_1, x_2, x_3, x_4) = k_2 x_3. \tag{2.26}$$

This problem was analyzed in [32], using involved arguments that were specific to the system. The signs of the partial derivatives of a concentration with respect to an initial concentration are given in Table 4. The * entries in Table 4 mean that the

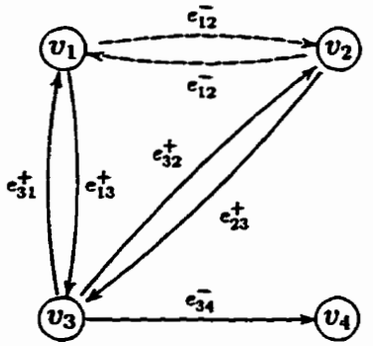


Figure 8: $G(\tilde{f}, \Omega)$ for Example 10

partial derivative is of both signs. The 0 entries mean that the partial derivative is identically zero. The Jacobian matrix for this system is

	$x_1(t)$	$x_2(t)$	$x_3(t)$	$x_4(t)$
$x_1(0)$	+	−	*	+
$x_2(0)$	−	+	+	+
$x_3(0)$	+	*	*	+
$x_4(0)$	0	0	0	+

Table 4: Behaviour of concentrations with respect to changes in initial concentrations for the Michaelis-Menten system for Example 10.

$$D\tilde{f} = \begin{pmatrix} -k_1x_2 & -k_1x_1 & k_{-1}+k_2 & 0 \\ -k_1x_2 & -k_1x_1 & k_{-1} & 0 \\ k_1x_2 & k_1x_1 & -(k_{-1}+k_2) & 0 \\ 0 & 0 & k_2 & 0 \end{pmatrix}.$$

Figure 8 presents the multigraph $G(\tilde{f}, \Omega)$ for the Michaelis-Menten system. Almost all ordered vertex pairs are inconsistently strongly connected in $G(\tilde{f}, \Omega)$; we can

only conclude that

$$\frac{\partial x_i}{\partial x_4(0)} = 0,\ i = 1,2,3,\ \text{and}\ \frac{\partial x_4}{\partial x_4(0)} > 0,\ t \geq 0.$$

Example 11 (Chemical Kinetics): We look at the complicated mechanism:

$$\begin{aligned}
A_1 + A_6 &\xrightarrow{k_1} A_1,\\
A_2 &\xrightarrow{k_2} A_4 + A_5,\\
A_2 + A_3 &\xrightarrow{k_3} A_2,\\
A_3 + A_6 &\xrightarrow{k_4} A_6,\\
A_3 &\xrightarrow{k_5} A_1,\\
A_6 &\xrightarrow{k_6} A_2,\\
A_4 &\xrightarrow{k_7} A_1 + A_6,\\
A_5 &\xrightarrow{k_8} A_6,\ \text{and}\\
A_3 + A_4 &\xrightarrow{k_9} A_4.
\end{aligned}$$

The corresponding system of differential equations is

$$\dot{x}_1 = k_5 x_3 + k_7 x_4, \tag{2.27}$$

$$\dot{x}_2 = -k_2 x_2 + k_6 x_6, \tag{2.28}$$

$$\dot{x}_3 = -k_3 x_2 x_3 - k_4 x_3 x_6 - k_5 x_3 - k_9 x_3 x_4, \tag{2.29}$$

$$\dot{x}_4 = k_2 x_2 - k_7 x_4, \tag{2.30}$$

$$\dot{x}_5 = k_2 x_2 - k_8 x_5,\ \text{and} \tag{2.31}$$

$$\dot{x}_6 = -k_1 x_1 x_6 - k_6 x_6 + k_7 x_4 + k_8 x_5, \tag{2.32}$$

and we construct the multigraph in Figure 9, where dashed edges have negative sign and solid edges have positive sign. One spanning tree of the associated graph

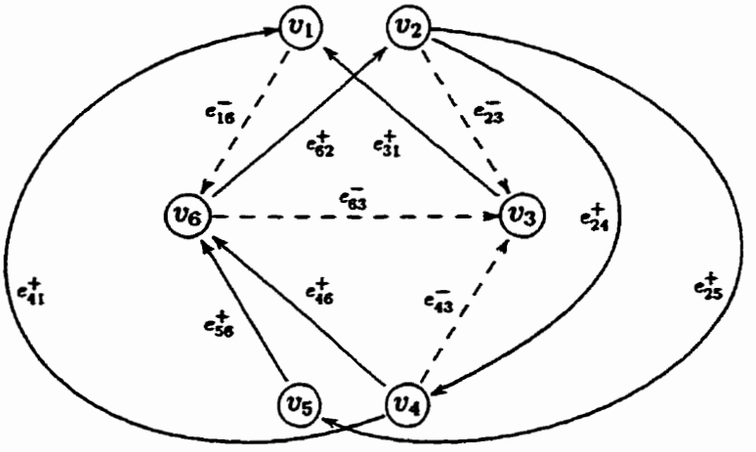


Figure 9: $G(\tilde{f}, \Omega)$ for Example 11

G_u and the fundamental set of cycles that it induces are given in Figure 10. As listed, the fifth cycle is negative, confirming that the mechanism does not induce an order preserving flow. In fact, performing the algorithm of Section 4 leads to a collapsed graph consisting of just one negative vertex; all vertices are inconsistently strongly connected.

Example 12 (Chemical Kinetics): Consider the reaction mechanism given by

$$\begin{aligned}
A_1 + A_6 &\xrightarrow{k_1} A_1, \\
A_2 &\xrightarrow{k_2} A_4 + A_5, \\
A_2 + A_3 &\xrightarrow{k_3} A_2, \\
A_3 + A_6 &\xrightarrow{k_4} A_6, \\
A_3 &\xrightarrow{k_5} A_1, \\
A_6 &\xrightarrow{k_6} A_2, \\
A_4 &\xrightarrow{k_7} A_6, \\
A_1 + A_4 &\xrightarrow{k_8} A_4, \\
A_5 &\xrightarrow{k_9} A_6, \text{ and}
\end{aligned}$$

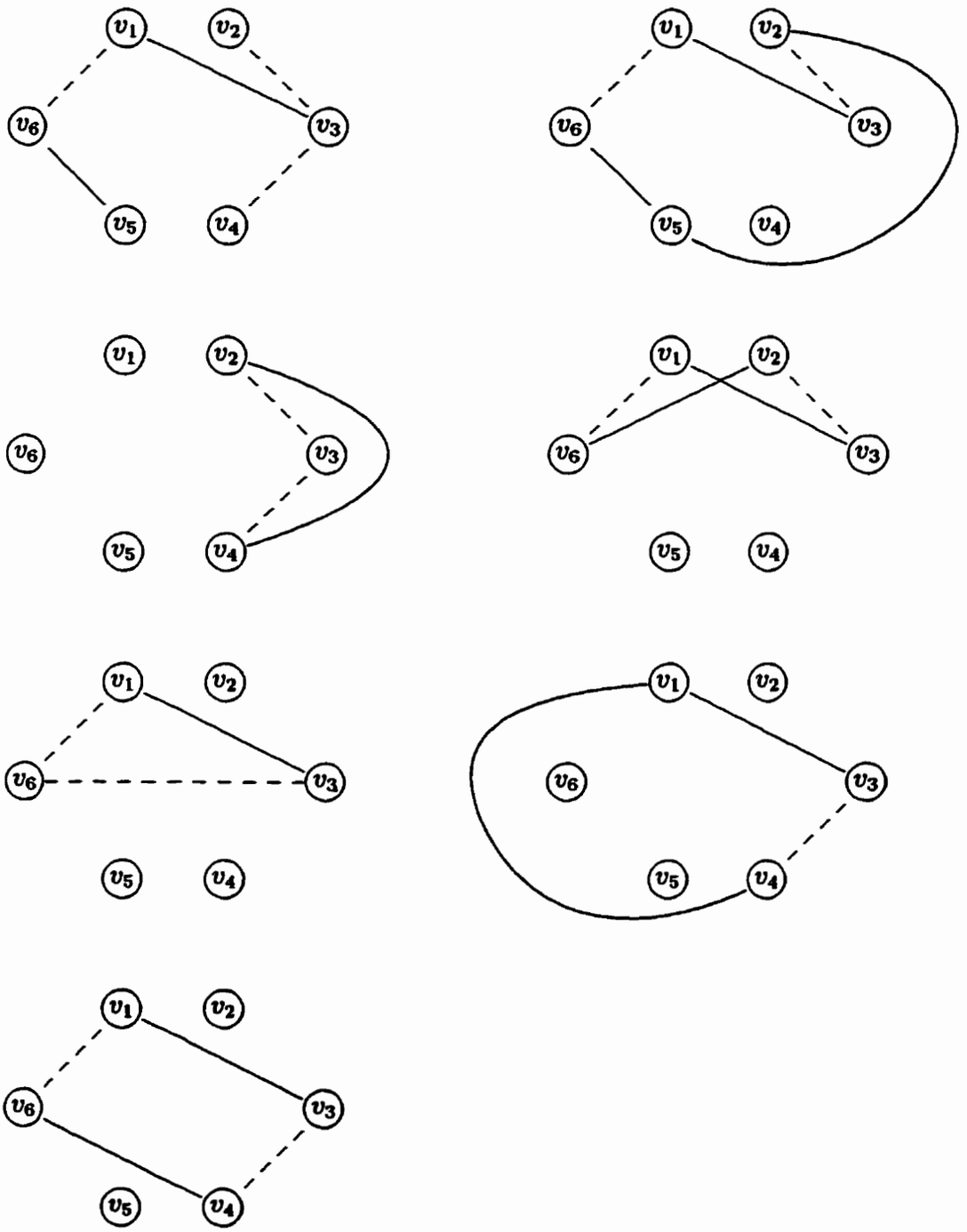


Figure 10: Spanning tree and fundamental set of cycles for Example 11.

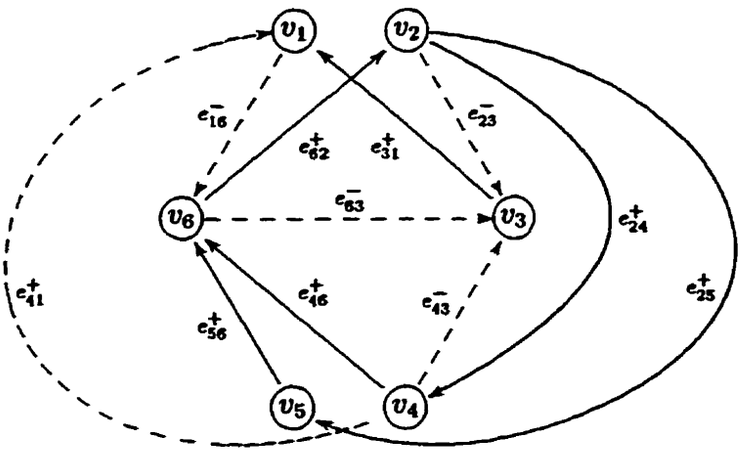


Figure 11: $G(\tilde{f}, \Omega)$ for Example 12

$$A_3 + A_4 \xrightarrow{k_{10}} A_4.$$

The corresponding multigraph is presented in Figure 11. This multigraph differs very little from the multigraph in previous example: the directed edge from v_4 to v_1 is now negative. Once again, we choose one spanning tree of the associated graph G_u and draw it and the fundamental set of cycles that it induces; see Figure 12. In this case, all of the cycles in the chosen fundamental set are positive; hence, by Theorem 5, the system induces an order preserving flow.

If the positivity assumption is satisfied, Theorem 8 applies and we can deduce the signs of the derivatives of the associated concentrations with respect to initial concentrations. These are presented in Table 5, where entries have the usual meaning.

The positivity assumption will be satisfied if $\tilde{x}_0 \in \mathcal{S}_0 = \bar{\mathcal{O}} \cap \{x_3 > 0$ and either $x_2 > 0$ or $x_4 > 0$ or $x_5 > 0$ or $x_6 > 0\}$. Using Theorem 18, we see that each reaction in this mechanism with a bimolecular reactant plays no role when analyzing positivity (every species occurs as a reactant or as a product in the reactions with a single reactant). The remaining reactions correspond to the positive edges in the

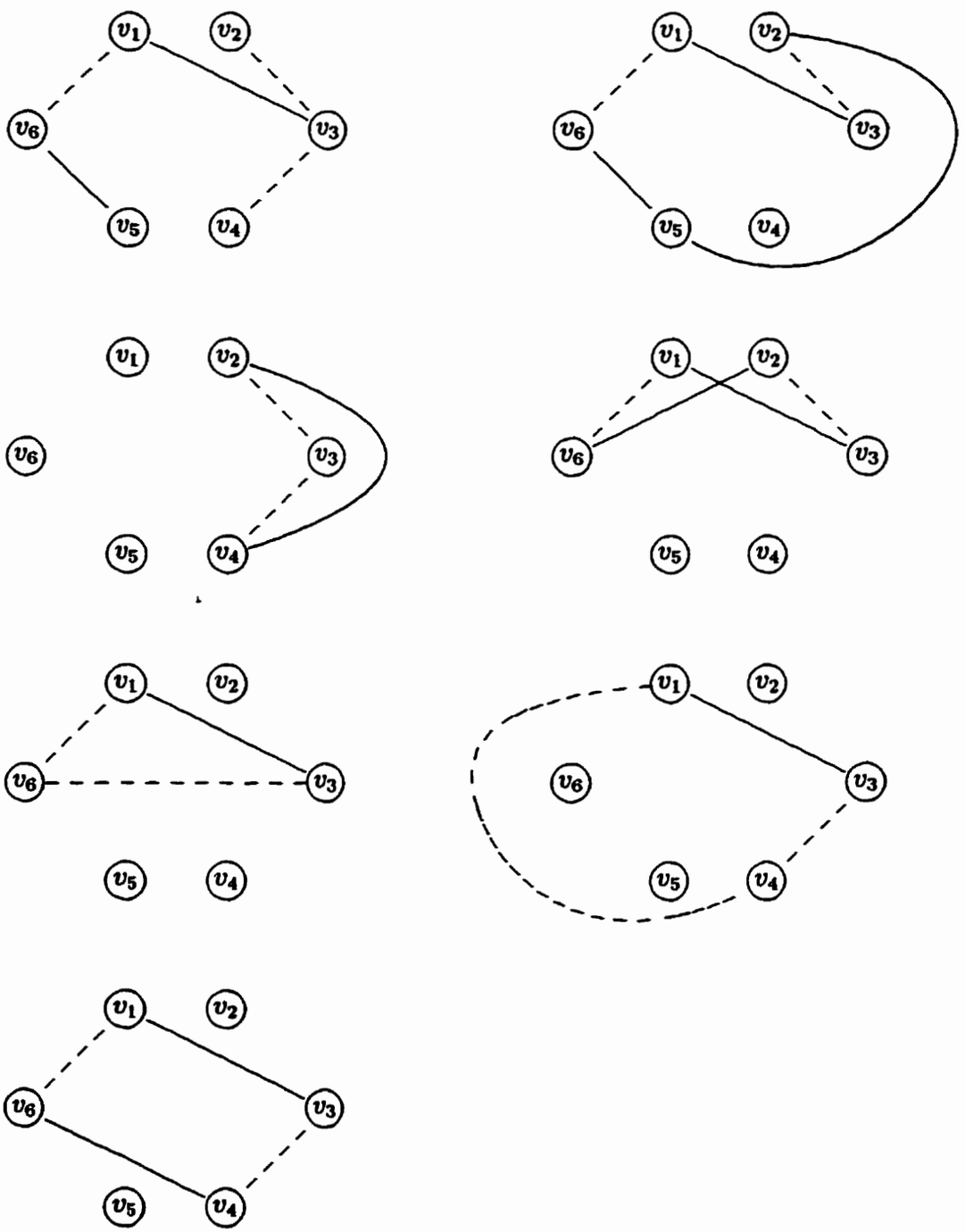


Figure 12: Spanning tree and fundamental set of cycles for Example 12.

	$x_1(t)$	$x_2(t)$	$x_3(t)$	$x_4(t)$	$x_5(t)$	$x_6(t)$
$x_{1,0}$	++	−	+	−	−	−
$x_{2,0}$	−	++	−	+	+	+
$x_{3,0}$	+	−	++	−	−	−
$x_{4,0}$	−	+	−	++	+	+
$x_{5,0}$	−	+	−	+	++	+
$x_{6,0}$	−	+	−	+	+	++

Table 5: Signs of concentrations with respect to changes in initial concentrations for Example 12.

graph of Figure 13. There are two initial groups: $\{v_3\}$ and $\{v_2, v_4, v_5, v_6\}$. The conclusion follows.

Example 13 (Chemical Kinetics): We look at the reaction mechanism:

$$\begin{aligned}
A_1 + A_6 &\xrightarrow{k_1} A_1, \\
A_2 &\xrightarrow{k_2} A_4 + A_5, \\
A_2 + A_3 &\xrightarrow{k_3} A_2, \\
A_3 + A_6 &\xrightarrow{k_4} A_6, \\
A_3 &\xrightarrow{k_5} A_1, \\
A_6 &\xrightarrow{k_6} A_2, \\
A_4 &\xrightarrow{k_7} A_6, \\
A_1 + A_4 &\xrightarrow{k_8} A_4, \\
A_5 &\xrightarrow{k_9} A_6, \\
A_3 + A_4 &\xrightarrow{k_{10}} A_4, \text{ and} \\
A_1 + A_2 + A_7 &\xrightarrow{k_{11}} A_1 + A_2.
\end{aligned}$$

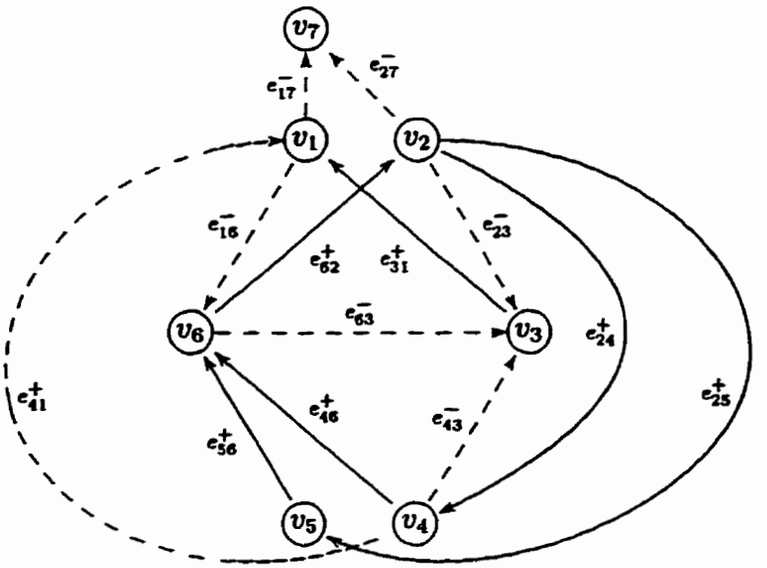


Figure 13: $G(\tilde{f}, \Omega)$ for Example 13

The corresponding system of differential equations is

$$\dot{x}_1 = -k_8 x_1 x_4 + k_5 x_3, \tag{2.33}$$

$$\dot{x}_2 = -k_2 x_2 + k_6 x_6, \tag{2.34}$$

$$\dot{x}_3 = -k_3 x_2 x_3 - k_4 x_6 x_3 - k_5 x_3 - k_{10} x_3 x_4, \tag{2.35}$$

$$\dot{x}_4 = k_2 x_2 - k_7 x_4, \tag{2.36}$$

$$\dot{x}_5 = k_2 x_2 - k_9 x_5, \tag{2.37}$$

$$\dot{x}_6 = -k_1 x_1 x_6 - k_6 x_6 + k_7 x_4 + k_9 x_5, \text{ and} \tag{2.38}$$

$$\dot{x}_7 = -k_{11} x_1 x_2 x_7, \tag{2.39}$$

which leads to the multigraph in Figure 13, where dashed edges have negative sign and solid edges have positive sign. Applying Lemma 14 tells us that vertices v_1 through v_6 in $G(\tilde{f}, \Omega)$ are each involved in at least one directed cycle. We observe that vertices v_1, v_6, v_2, and v_3 comprise a positive directed cycle. We collapse these four vertices into a positive vertex labelled $v_{1,6,2,3}$ and pick vertex v_1 to determine the signs of the edges in our collapsed graph. For example, there is a negative

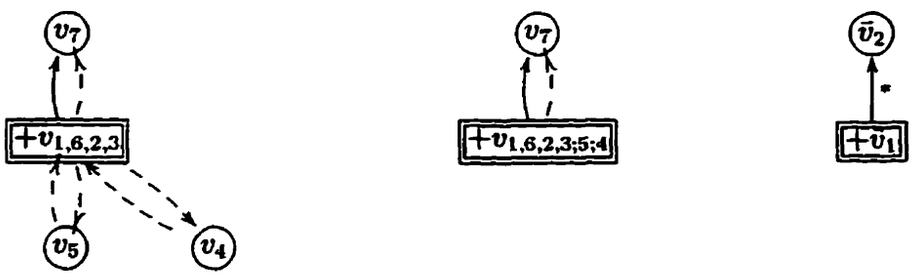


Figure 14: $G_1(\tilde{f},\Omega)$, $G_3(\tilde{f},\Omega)$, and $G_w(\tilde{f},\Omega)$ for Example 13.

edge from vertex v_4 to vertex v_3 and there is a positive edge from vertex v_3 to v_1; hence, we will draw a negative edge from v_4 to $v_{1,6,2,3}$ in the collapsed graph. Proceeding in this way leads to the leftmost multigraph in Figure 14. Vertices $v_{1,6,2,3}$ and v_5 obviously form a positive directed cycle; the same is true of vertices $v_{1,6,2,3}$ and v_4. We can collapse both of these cycles to get the middle multigraph in Figure 14. Finally, $G_w(\tilde{f},\Omega)$ is constructed on the right in Figure 14. Here, $n_w = 2$ and we need only observe that $w_{12} = *$. We conclude that vertices v_1 through v_6 are consistently strongly connected in $G(\tilde{f},\Omega)$ and can deduce the signs of the derivatives of the associated concentrations with respect to initial concentrations. These are presented in Table 6, where entries have the usual meaning.

	$x_1(t)$	$x_2(t)$	$x_3(t)$	$x_4(t)$	$x_5(t)$	$x_6(t)$	$x_7(t)$
$x_{1,0}$	++	−	+	−	−	−	
$x_{2,0}$	−	++	−	+	+	+	
$x_{3,0}$	+	−	++	−	−	−	
$x_{4,0}$	−	+	−	++	+	+	
$x_{5,0}$	−	+	−	+	++	+	
$x_{6,0}$	−	+	−	+	+	++	
$x_{7,0}$	0	0	0	0	0	0	++

Table 6: Signs of concentrations with respect to changes in initial concentrations for Example 13.

Chapter 3

Monotonicity With Respect to a Cone

In previous work (see [30], [32], [18], [31]), partial monotonicity results were proven for several specific chemical and epidemiological models. Solutions were only monotone with respect to initial conditions in a subset of solution space.

The treatment of linear systems in [5] and [4] combined with our experience with some specific nonlinear problems motivates an attempt to formulate monotonicity results with respect to a convex cone. In the next section, we provide an extension of the Kamke-Müller Theorem to closed, convex cones.

3.1 Preliminaries

Firstly, we need to define the concept of a cone in $\mathbf{R}^n$.

Definition 5 *A set $K \subseteq \mathbf{R}^n$ is defined to be a* **cone** *if , $\forall\ \tilde{x} \in K$ and $\alpha \geq 0$, $\alpha\tilde{x} \in K$. A cone K is said to be* **solid** *if $int(K) \neq \emptyset$. K is said to be* **pointed** *if*

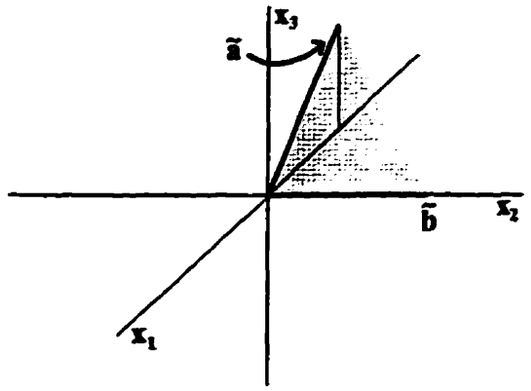


Figure 15: A two-dimensional cone K in $\mathbf{R}^3$ is not solid.

$K \cap \{-K\} = \{\tilde{0}\}$.

One must be careful when discussing non-solid cones. For example, consider the two-dimensional cone K in $\mathbf{R}^3$ presented in Figure 15. K is not solid: for any point in K, an arbitrarily small ball in $\mathbf{R}^3$ around that point is not in K. However, K is solid when we restrict ourselves to the smallest subspace containing it (the plane spanned by $\tilde{a}$ and $\tilde{b}$ in the example). The cone's interior relative to the smallest subspace containing K is called the *relative interior* of K; we denote it *relint*(K). Similarly, the cone's boundary relative to the smallest subspace containing K is called the *relative boundary* of K and is denoted *relbdy*(K).

Recall that a set in $\mathbf{R}^n$ is convex if for any two of its points it contains the line segment between them. A convex cone K induces a partial ordering "$\leq_K$" in $\mathbf{R}^n$. For $\tilde{x}, \tilde{y} \in \mathbf{R}^n$, we write $\tilde{x} \leq_K \tilde{y}$ (or $\tilde{y} \geq_K \tilde{x}$) if and only if $\tilde{y} - \tilde{x} \in K$. We will write $\tilde{x} <_K \tilde{y}$ (or $\tilde{y} >_K \tilde{x}$) if $\tilde{y} - \tilde{x} \in$ *relint*(K). The partial ordering induced by a convex cone K is antisymmetric ($\tilde{v} \geq_K \tilde{w}$ and $\tilde{w} \geq_K \tilde{v} \Rightarrow \tilde{v} = \tilde{w}$) if and only if K is pointed.

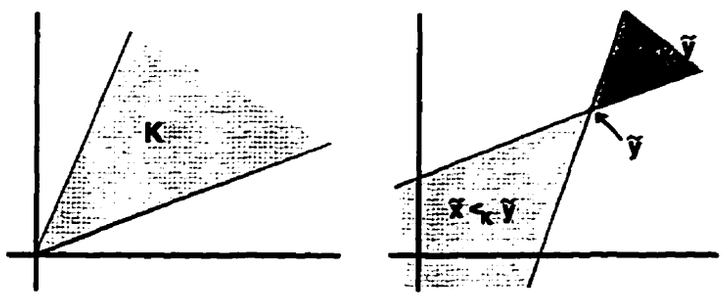


Figure 16: Orderings with respect to a proper cone K in $\mathbf{R}^2$

A *proper cone* is a cone which is closed, convex, solid and pointed. A proper cone is generated by its extreme rays (all vectors in the cone are a non-negative, linear combination of the extreme rays; see page three of [5]). A vector $\tilde{x}$ is an *extreme ray* of K if $\tilde{0} \leq_K \tilde{y} \leq_K \tilde{x} \Rightarrow \tilde{y}$ is a non-negative multiple of $\tilde{x}$. The early pages of [5] and [43] offer an introduction to this terminology and theory.

Geometrically, in $\mathbf{R}^2$, Figure 16 illustrates which vectors $\tilde{x}$ satisfy the two strict inequalities for a fixed vector $\tilde{y}$ and a proper cone K. As it turns out, we will want to verify that a chosen cone satisfies particular hypotheses in order to apply the upcoming results. Since proper cones are generated by their extreme rays, we need only check that the extreme rays satisfy the properties we demand of all rays in the cone.

All seems well, but we will frequently use *polyhedral cones* in practice. The cone in Figure 16 is a *polyhedral cone* as well as a proper cone.

[5] offers the following comments on polyhedral cones (Theorem 2.5, page 2):

(1) A nonempty set K of $\mathbf{R}^n$ is a polyhedral cone if and only if it is the intersection of a finitely many closed half spaces, each containing the origin on its boundary;

(2) A polyhedral cone is a closed, convex cone; and

(3) A nonempty subset K of $\mathbb{R}^n$ is a polyhedral cone if and only if K^* is a polyhedral cone.

K^* denotes the *dual cone* of K, given by

$$K^* = \{\tilde{y} \in \mathbb{R}^n : \tilde{x} \in K \Rightarrow \tilde{x} \cdot \tilde{y} \geq 0\}.$$

Note that $K^{**} = K$ if and only if K is a closed, convex cone (see [5]).

The right, circular ("ice-cream") cone with vertex at the origin is a proper cone that is not polyhedral. It would seem that the defining difference is that proper cones do not necessarily have a finite number of extreme rays (or generators). Note that we will use the word generators here because the notion of extreme rays will not make sense for some polyhedral cones we consider. It is not quite as simple as one might hope to establish that polyhedral cones do have a finite number of generators. We will need a small amount of theory.

We need to define the concept of a face. Let K and $F \subseteq K$ be pointed, closed cones; then F is called a *face* of K if

$$\tilde{x} \in F \text{ and } \tilde{0} \leq_K \tilde{y} \leq_K \tilde{x} \Rightarrow \tilde{y} \in F.$$

The face F is nontrivial if $F \neq \{\tilde{0}\}$ and $F \neq K$. For example, the faces of the non-negative orthant $\bar{\mathcal{O}}$ are of the form $F_J = \{\tilde{x} \in \bar{\mathcal{O}} : x_j = 0 \text{ if } j \notin J\}$ where $J \subseteq \{1, \ldots, n\}$. This includes the two-dimensional faces that one might think of naturally, along with the one-dimensional faces (the extreme rays of $\bar{\mathcal{O}}$) and the trivial faces ($\tilde{0}$ and $\bar{\mathcal{O}}$). As a second example, note that the nontrivial faces of the ice-cream cone with vertex at the origin are of the form $\alpha\tilde{x}$ where $\alpha > 0$ and $\tilde{x}$ is a boundary vector. The non-negative orthant has a finite number of extreme rays (and faces) while the right circular cone has an infinite number.

Corollary 2.6.14 in [42] gives some insight into how the finite number of generators are chosen: *Every closed, convex set in* $\mathbb{R}^n$ *is the convex hull of those of its faces which are flats or closed halfflats.* A set $A \in \mathbb{R}^n$ is called a *flat* if whenever it contains two points, it also contains the entire line through them. A *closed halfflat* is the intersection of a flat with a closed halfspace which meets it, but does not contain it. We now present the result from [42] (Theorem 4.1.1) that satisfies our needs.

Theorem 19 *A convex cone in* $\mathbb{R}^n$ *is finitely generated if and only if it is polyhedral.*

Consider the polyhedral cone K in Figure 17; K is the cone in $\mathbb{R}^3$ defined by $\{(x, y, z) : z \geq |y|\}$. K is polyhedral (it is the intersection of two halfspaces, each with the origin on its boundary); but, it is not proper (it is not pointed). Its faces are two closed halfflats (the two halfplanes) and one flat (the x-axis). Following the results above, its generators are

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix},$$

where the last two vectors could each be chosen to be any vector in each of the halfplanes but not on the x-axis. Notice that neither of the final two vectors, however chosen, are extreme vectors.

The following results will be useful.

Lemma 20 *Let* S *be a convex subset of* $\mathbb{R}^n$ *and let* $\tilde{x} \in relbdy(S)$. *Then there exists a hyperplane* H *containing* $\tilde{x}$ *such that* S *is contained in one of the halfspaces associated with* H. *We call* H *a supporting hyperplane at* $\tilde{x}$. *The vector* $\tilde{b} \in \mathbb{R}^n$ *is*

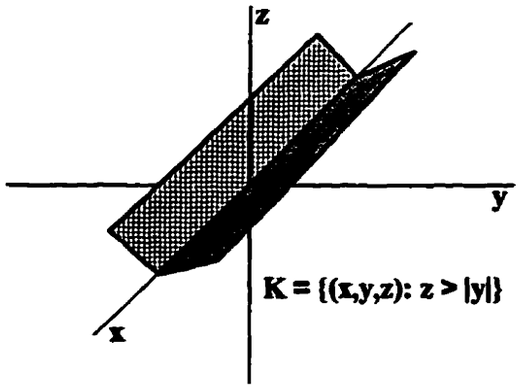


Figure 17: An unpointed polyhedral cone in $\mathbb{R}^3$

a normal to H if $\tilde{b} \neq \tilde{0}$ is orthogonal to the difference of any two vectors in H. We say $\tilde{b}$ is normal to S at $\tilde{x}$. Furthermore, if $\tilde{b}$ satisfies $\tilde{b} \cdot (\tilde{y} - \tilde{x}) \leq 0$, $\forall\ \tilde{x}, \tilde{y} \in S$, we say that $\tilde{b}$ is an outward normal to S at $\tilde{x}$.

For a proof of this result, see Theorem 2.7 in [4]. Books on linear programming (such as [10]) also introduce this result and discuss convex cones.

Lemma 21 *Let X be a compact subset of $\mathbb{R}^n$ and K be a closed, convex cone in $\mathbb{R}^n$ with $\tilde{k} \in relint(K)$. Then there is a positive constant α such that $\alpha\tilde{k} + \tilde{x} >_K \tilde{0}$ for all $\tilde{x} \in X$.*

Proof: It is sufficient to show that there exists a positive α so that $\tilde{k} + \frac{\tilde{x}}{\alpha} \in relint(K)$ for all $\tilde{x} \in X$. Since $\tilde{k}$ is a relative interior vector of the closed, convex cone K, there is a ball of radius ϵ, for some $\epsilon > 0$, centered at $\tilde{k}$, contained in $relint(K)$. Denote this ball by $B_\epsilon(\tilde{k})$. Then $B_\epsilon(\tilde{k}) = B_\epsilon(\tilde{0}) + \tilde{k} \subset relint(K)$. Let $d = max\{|\tilde{x}| : \tilde{x} \in X\}$ and choose $\alpha > d$. Then $\frac{\tilde{x}}{\alpha} \in B_\epsilon(\tilde{0})$ for all $\tilde{x} \in X$, and the conclusion follows. □

Since we will want to allow non-solid cones in applications, we need the following result. It guarantees that a difference of super- and sub-solutions which begins in

a closed, convex cone K stay in the smallest subspace containing K under the hypotheses of the key results in this chapter.

Lemma 22 *For K a closed, convex cone in $\mathbb{R}^n$, suppose $\dot{\tilde{v}} \leq_K \tilde{f}(\tilde{v})$, $\dot{\tilde{w}} \geq_K \tilde{f}(\tilde{w})$, and $\tilde{w}(0) \geq_K \tilde{v}(0)$. Suppose also that $\tilde{f}(\tilde{x})$ is continuously differentiable in $\tilde{x}$ on compact subsets of $\mathbb{R}^n$ and that for any compact set, N, $\exists\ l = l(N)$ such that*

$$D\tilde{f}(\tilde{x}) + lI : K \mapsto K, \forall\ \tilde{x} \in N;$$

then $\tilde{w} - \tilde{v} \in \pi$, where π is the smallest subspace containing K, for $t \geq 0$.

Proof: Rewrite the hypotheses, $\dot{\tilde{w}} - \tilde{f}(\tilde{w}) \in K$ and $\tilde{f}(\tilde{v}) - \dot{\tilde{v}} \in K$, giving

$$\dot{\tilde{w}} - \dot{\tilde{v}} + \tilde{f}(\tilde{v}) - \tilde{f}(\tilde{w}) \in K. \tag{3.1}$$

Now, $\forall\ \tilde{w}$ in some compact set N, $\tilde{k} \in K$, $\tilde{k}^* \in K^*$,

$$\left[(D\tilde{f}(\tilde{w} + s\tilde{k}))\tilde{k} + l\tilde{k}\right] \cdot \tilde{k}^* \geq 0, \tag{3.2}$$

where we have strategically chosen the argument of $D\tilde{f}$ and $s \in [0, 1]$. Notice that if $\tilde{g}(s) = \tilde{F}(\tilde{w} + s\tilde{k})$, then $\tilde{g}'(s) = D\tilde{F}(\tilde{w} + s\tilde{k})\tilde{k}$; hence, if we define

$$g(s) = \left[\tilde{f}(\tilde{w} + s\tilde{k}) + sl\tilde{k}\right] \cdot \tilde{k}^*,$$

then (3.2) says that $g'(s) \geq 0$. We can conclude that $g(1) \geq g(0)$, say. This gives

$$\left[\tilde{f}(\tilde{w} + \tilde{k}) + l\tilde{k} - \tilde{f}(\tilde{w})\right] \cdot \tilde{k}^* \geq 0,$$

which, upon letting $\tilde{v} = \tilde{w} + \tilde{k} \geq_K \tilde{w}$, leads to

$$\tilde{f}(\tilde{v}) - \tilde{f}(\tilde{w}) + l(\tilde{v} - \tilde{w}) \geq_K \tilde{0},$$

or

$$\tilde{f}(\tilde{v}) - \tilde{f}(\tilde{w}) + l(\tilde{v} - \tilde{w}) = \tilde{a},$$

where $\tilde{a} \in K$. So, (3.1) becomes

$$\dot{\tilde{w}} - \dot{\tilde{v}} + l(\tilde{w} - \tilde{v}) + \tilde{a} = \tilde{a}_1 \in K.$$

Let $h(t) = \tilde{b} \cdot (\tilde{w} - \tilde{v})$, where $\tilde{b}$ is normal to the subspace π; then,

$$\begin{aligned} \dot{h}(t) &= \tilde{b} \cdot (\dot{\tilde{w}} - \dot{\tilde{v}}) \\ &= -l\tilde{b} \cdot (\tilde{w} - \tilde{v}) + \tilde{b} \cdot (\tilde{a}_1 - \tilde{a}) \\ &= -l\tilde{b} \cdot (\tilde{w} - \tilde{v}) \qquad \text{since } \tilde{a}_1 - \tilde{a} \in K \\ &= -lh(t) \end{aligned}$$

and $h(0) = 0$, since $\tilde{w}(0) - \tilde{v}(0) \in K$. This implies that $h(t) = 0$, $\forall\, t \geq 0$, and the conclusion follows. □

We are now able to present an extension of the Kamke-Müller Theorem to closed, convex cones. The results of this section will generally be stated "$\forall\, t > 0$," as opposed to for "$0 < t < T$" or "$t \in I(\tilde{x})$" as in the previous chapter. We are assuming that solutions exist for all time; this will be the case in our examples. The results could be reformulated if one wanted to deal with intervals of existence.

Theorem 23 (Extended Kamke-Müller Theorem) *For K a closed, convex cone in $\mathbf{R}^n$, suppose $\dot{\tilde{v}} \leq_K \tilde{f}(\tilde{v})$, $\dot{\tilde{w}} \geq_K \tilde{f}(\tilde{w})$, and $\tilde{w}(0) \geq_K \tilde{v}(0)$. Suppose also that $\tilde{f}(\tilde{x})$ is continuously differentiable in $\tilde{x}$ on compact subsets of Ω, Ω open and convex, and that for any compact set, N, $\exists\, l = l(N)$ such that*

$$D\tilde{f}(\tilde{x}) + lI : K \mapsto K, \forall\, \tilde{x} \in N.$$

Then $\tilde{w}(t) \geq_K \tilde{v}(t)$, $\forall\, t \geq 0$.

Proof: Since $\tilde{w}(0) - \tilde{v}(0) \in K$, by Lemma 22, $\tilde{w}(t) - \tilde{v}(t) \in \pi$ for $t > 0$, where π is the smallest subspace of $\mathbf{R}^n$ containing K. We first suppose that $\dot{\tilde{v}} \leq_K \tilde{f}(\tilde{v})$,

$\dot{\tilde{w}} >_K \tilde{f}(\tilde{w})$, and $\tilde{w}(0) >_K \tilde{v}(0)$. We will prove that $\tilde{w}(t) >_K \tilde{v}(t)$, $\forall\ t \geq 0$. Suppose this is the not the case. Then $\exists\ t_0 > 0$ such that $\tilde{w}(t) >_K \tilde{v}(t)$ for $0 \leq t < t_0$ and $\tilde{w}(t_0) - \tilde{v}(t_0) \in relbdy(K)$. Let N be a compact, convex set containing both trajectories $\tilde{w}(t)$ and $\tilde{v}(t)$ for $0 \leq t < t_0$ (choose a large enough closed ball containing both). Let $\tilde{z}(t) = \tilde{w}(t) - \tilde{v}(t)$. If $\dot{\tilde{z}}(t_0) \in relint(K)$ then $\exists\ t_1$, $0 \leq t_1 < t_0$, such that $\dot{\tilde{z}}(t) \in relint(K)$ for $t_1 < t \leq t_0$. Then $\tilde{z}(t) \in relint(K)$ for $0 < t < t_0$ and $\dot{\tilde{z}}(t) \in relint(K)$ for $t_1 < t \leq t_0$ give $\tilde{z}(t_0) \in relint(K)$, a contradiction. Hence, $\dot{\tilde{z}}(t_0) \notin relint(K)$. By Lemma 20, $\exists$ a supporting hyperplane to K at $\tilde{z}(t_0)$. Let $\tilde{b}$ be the outward unit normal to K at $\tilde{z}(t_0)$. Since $\tilde{z}(t_0) \in relbdy(K)$ and $\dot{\tilde{z}}(t_0) \notin relint(K)$ we have $\tilde{z}(t_0) \cdot \tilde{b} = 0$ and $\dot{\tilde{z}}(t_0) \cdot \tilde{b} \geq 0$. Now, by the hypotheses, $\dot{\tilde{z}}(t) = \dot{\tilde{w}}(t) - \dot{\tilde{v}}(t) >_K \tilde{f}(\tilde{w}) - \tilde{f}(\tilde{v})$, $\forall\ t \geq 0$. Evaluating at t_0 gives $\dot{\tilde{z}}(t_0) - \tilde{f}(\tilde{w}(t_0)) + \tilde{f}(\tilde{v}(t_0)) >_K \tilde{0}$. Thus,

$$\left[\dot{\tilde{z}}(t_0) - \tilde{f}(\tilde{w}(t_0)) + \tilde{f}(\tilde{v}(t_0))\right] \cdot \tilde{b} < 0 \ \Rightarrow\ \left[\tilde{f}(\tilde{w}(t_0)) - \tilde{f}(\tilde{v}(t_0))\right] \cdot \tilde{b} > \dot{\tilde{z}}(t_0) \cdot \tilde{b} \geq 0$$
$$\Rightarrow\ \left[\tilde{f}(\tilde{w}(t_0)) - \tilde{f}(\tilde{v}(t_0))\right] \notin K$$

Now,

$$\tilde{f}(\tilde{w}(t_0)) - \tilde{f}(\tilde{v}(t_0)) = \left[\int_0^1 D\tilde{f}(s\tilde{w}(t_0) + (1-s)\tilde{v}(t_0))ds\right] \tilde{z}(t_0).$$

Since, by assumption and by the closure of K,

$$\left[\int_0^1 D\tilde{f}(s\tilde{w}(t_0) + (1-s)\tilde{v}(t_0))ds + lI\right] \tilde{z}(t_0) \in K,$$

we have

$$\left[\int_0^1 D\tilde{f}(s\tilde{w}(t_0) + (1-s)\tilde{v}(t_0))ds + lI\right] \tilde{z}(t_0) \cdot \tilde{b} \leq 0.$$

But

$$\left[\int_0^1 D\tilde{f}(s\tilde{w}(t_0) + (1-s)\tilde{v}(t_0))ds + lI\right] \tilde{z}(t_0) \cdot \tilde{b}$$
$$= \left[\int_0^1 D\tilde{f}(s\tilde{w}(t_0) + (1-s)\tilde{v}(t_0))ds\right] \tilde{z}(t_0) \cdot \tilde{b}, \qquad \text{since } \tilde{z}(t_0) \cdot \tilde{b} = 0$$

$$
\begin{aligned}
&= \left[\tilde{f}(\tilde{w}(t_0)) - \tilde{f}(\tilde{v}(t_0))\right] \cdot \tilde{b} \\
&> 0,
\end{aligned}
$$

gives a contradiction; hence, $\tilde{w}(t) >_K \tilde{v}(t)$, $\forall\, t \geq 0$.

To prove the theorem, we let $\tilde{w}^\epsilon(t) = \tilde{w}(t) + \epsilon e^{\alpha t}\tilde{k}$, where ϵ is a small positive parameter, $\tilde{k}$ is a relative interior vector of K, and α is a constant to be chosen later. Then

$$\tilde{w}^\epsilon(0) = \tilde{w}(0) + \epsilon\tilde{k} >_K \tilde{w}(0) \geq_K \tilde{v}(0).$$

Now, let X be the compact set $\{\tilde{f}(\tilde{w}(t)) - \tilde{f}(\tilde{w}^\epsilon(t)) : t \in [0,T]\}$. Lemma 21 tells us that for some $\beta > 0$ and for all $t \in [0,T]$

$$\tilde{f}(\tilde{w}(t)) - \tilde{f}(\tilde{w}^\epsilon(t)) + \beta\tilde{k} >_K 0.$$

Now, choosing $\alpha = \frac{\beta}{\epsilon}$, we have for $t \in [0,T]$

$$\dot{\tilde{w}}^\epsilon(t) = \dot{\tilde{w}}(t) + \epsilon\alpha e^{\alpha t}\tilde{k} \geq_K \tilde{f}(\tilde{w}(t)) + \beta\tilde{k} >_K \tilde{f}(\tilde{w}^\epsilon(t)).$$

By the first result in this proof, we can conclude that $\tilde{v}(t) <_K \tilde{w}^\epsilon(t) = \tilde{w}(t) + \epsilon e^{\alpha t}\tilde{k}$, for $t \in [0,T]$. Letting $\epsilon \to 0$ proves the theorem. □

It turns out that there is some development of this type of result in the literature (see [40], [37], [38], [28], and [29]). In particular, [29] presents a result very similar to the above, but the condition

$$\exists l = l(N) \text{ such that } D\tilde{f}(\tilde{x}) + lI : K \mapsto K, \forall\, \tilde{x} \in N, \tag{3.3}$$

is not given. Later in this chapter, we will show how the various conditions in the literature and condition (3.3) are linked. When (3.3) holds, we will say that $D\tilde{f}(\tilde{x})$ preserves the cone K. A closed, convex polyhedral cone is determined by

its generators. Hence, we need only investigate the effect of $D\tilde{f}(\tilde{x}) + lI$ on the generators.

One might imagine the Extended Kamke-Müller Theorem being applied to two solutions of a system of ordinary differential equations, one with a single component changed initially. When the hypotheses are satisfied, we could conclude that the two solution vectors maintain a partial ordering with respect to the cone used (finding useful cones is a thorny issue we will deal with later). Still, we would like to be able to draw conclusions on partial derivatives with respect to initial conditions. So, suppose we are considering the usual system of ordinary differential equations (2.7) and that we have found a cone K which satisfies condition (3.3). Apply Theorem 23 with $\tilde{v}(t) = \tilde{x}(t, \tilde{x}_0)$ and $\tilde{w}(t) = \tilde{x}(t, \tilde{x}_0 + \tilde{\beta})$, with $\tilde{\beta} \in K \setminus \{\tilde{0}\}$, and $\tilde{x}(t, \tilde{x}_0)$ the solution to (2.7) with initial condition $\tilde{x}(0) = \tilde{x}_0$, $\tilde{x}_0 \in \Omega$, to conclude that $\tilde{w}(t) \geq_K \tilde{v}(t)$. In other words, we conclude that

$$\tilde{x}(t, \tilde{x}_0 + \tilde{\beta}) - \tilde{x}(t, \tilde{x}_0) \geq_K \tilde{0}.$$

Pick $\tilde{\beta} = s\tilde{u} \in K \setminus \{\tilde{0}\}$, $s > 0$, where $\tilde{u}$ is a unit vector; then

$$D_{\tilde{u}}\tilde{x}(t) = \lim_{s \to 0} \frac{\tilde{x}(t, \tilde{x}_0 + s\tilde{u}) - \tilde{x}(t, \tilde{x}_0)}{s} = \lim_{s \to 0^+} \frac{\tilde{x}(t, \tilde{x}_0 + s\tilde{u}) - \tilde{x}(t, \tilde{x}_0)}{s}$$

represents the directional derivative of $\tilde{x}$ in the direction $\tilde{u}$. We have shown that $D_{\tilde{u}}\tilde{x} \in K$. When $\tilde{u}$ is a standard basis vector, this gives derivatives with respect to an initial component as before. Upcoming examples will illustrate the idea. We state the following Corollary for use in applications.

Corollary 24 *Suppose that $\tilde{\mathcal{O}}$ is positively invariant and that $\tilde{f}(\tilde{x})$ is continuously differentiable on $\tilde{\mathcal{O}}$ and that $\exists\ l$ such that*

$$D\tilde{f}(\tilde{x}) + lI : K \mapsto K, \forall\ \tilde{x} \in \tilde{\mathcal{O}},$$

where K is a closed, convex cone in $\mathbb{R}^n$; then

$$\frac{\partial \tilde{x}}{\partial \tilde{k}} \geq_K 0,$$

for any $\tilde{k} \in K \setminus \{\tilde{0}\}$,$\forall\, t \geq 0$.

Proof: Define the sequences $\{\tilde{x}_0^m\}$ and $\{\tilde{y}_0^m\}$ with $\tilde{y}_0^m \geq_K \tilde{x}_0^m$, $\tilde{x}_0^m, \tilde{y}_0^m \in \mathcal{O}$, $\lim_{m\to\infty} \tilde{y}_0^m = \tilde{y}_0^*$, and $\lim_{m\to\infty} \tilde{x}_0^m = \tilde{x}_0^*$, with $\tilde{x}_0^* \in bdy(\bar{\mathcal{O}})$. By Theorem 23, $\tilde{x}(t, \tilde{y}_0^m) \geq_K \tilde{x}(t, \tilde{x}_0^m)$. Using the limit property, $\tilde{x}(t, \tilde{y}_0^*) \geq_K \tilde{x}(t, \tilde{x}_0^*)$. Now choose $\tilde{y}_0^m = \tilde{x}_0^m + \tilde{k}$, $\tilde{k} \in K \setminus \{\tilde{0}\}$; then $\tilde{x}(t, \tilde{x}_0^* + \tilde{k}) \geq_K \tilde{x}(t, \tilde{x}_0^*)$, which leads to

$$\frac{\partial \tilde{x}}{\partial \tilde{k}} \geq_K \tilde{0}.$$

□

3.2 Practical Tools for Establishing Monotonicity

For preliminary investigation, and to highlight the complexity of this approach, we consider a 2×2 constant matrix,

$$\mathcal{M} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix},$$

and a general closed, convex cone with extreme rays $\tilde{a}$ and $\tilde{b}$, oriented as in Figure 18. In order for (3.3) to hold, with $\mathcal{M}$ replacing $D\tilde{f}(\tilde{x})$, $\mathcal{M}\tilde{a}$ must lie on the same side of the line containing $\tilde{a}$ as does $\tilde{a}^\perp$, the vector that is rotated 90° counterclockwise from $\tilde{a}$. Similarly, $\mathcal{M}\tilde{b}$ and $-\tilde{b}^\perp$ must lie on the same side of the line containing $\tilde{b}$. That is,

$$\mathcal{M}\tilde{a} \cdot \tilde{a}^\perp = \mathcal{M}\tilde{a} \cdot (-a_2, a_1)$$

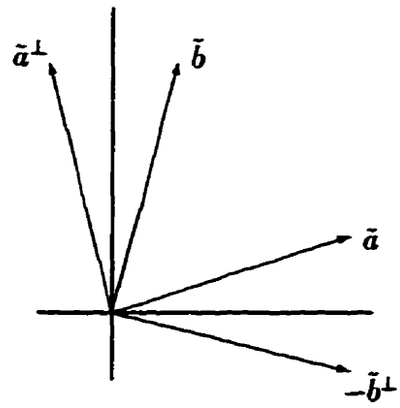


Figure 18: A 2-dimensional cone with extreme rays $\tilde{a}$ and $\tilde{b}$.

$$= m_{21}a_1^2 + a_1a_2(m_{22} - m_{11}) - m_{12}a_2^2 \geq 0, \text{ and} \tag{3.4}$$

$$\begin{aligned} \mathcal{M}\tilde{b} \cdot (-\tilde{b}^{\perp}) &= \mathcal{M}\tilde{b} \cdot (b_2, -b_1) \\ &= -m_{21}b_1^2 - b_1b_2(m_{22} - m_{11}) + m_{12}b_2^2 \geq 0. \end{aligned} \tag{3.5}$$

These are conditions on the quadratic form $Q(a_1, a_2) = \mathcal{M}\tilde{a} \cdot \tilde{a}^{\perp}$. As outlined in [44], through a rotation of the coordinate axes the mixed term in (3.4) and (3.5) will be eliminated when the form is expressed in terms of the new coordinates. In these new coordinates, the form can be expressed as

$$Q'(x_1, x_2) = \lambda_1 x_1^2 + \lambda_2 x_2^2, \tag{3.6}$$

where λ_1 and λ_2 are in fact the eigenvalues of the matrix

$$\tilde{\mathcal{M}} = \begin{bmatrix} m_{21} & m_{22} \\ -m_{11} & -m_{12} \end{bmatrix},$$

$\tilde{a}$ is rotated to a new vector $\tilde{a}'$, and $\tilde{b}$ is rotated to a new vector $\tilde{b}'$. With the form expressed as in (3.6), we require $Q'(\tilde{a}') \geq 0$ and $Q'(\tilde{b}') \leq 0$. This is possible if and only if $\lambda_1\lambda_2 \leq 0$; in this case, the form is called negative semidefinite.

A quadratic form $Ay_1^2+By_1y_2+Cy_2^2$ is negative semidefinite when $4AC-B^2 \leq 0$; thus, we have arrived at a necessary and sufficient condition for there to exist a cone preserved by the matrix $\mathcal{M}$, namely that

$$(m_{22}-m_{11})^2+4m_{12}m_{21} \geq 0, \text{ or equivalently} \tag{3.7}$$

$$tr(\mathcal{M})^2-4det(\mathcal{M}) \geq 0. \tag{3.8}$$

In the case $\lambda_1\lambda_2 < 0$, $Q'(x_1,x_2) = 0$ along exactly two lines, those with slopes $\pm\sqrt{-\lambda_1\lambda_2^{-1}}$ in the rotated coordinates, or $Q(a_1,a_2) = 0$ along lines with slope

$$\mu_1 = \frac{m_{22}-m_{11}+\sqrt{(m_{22}-m_{11})^2+4m_{12}m_{21}}}{2m_{12}}, \text{ and} \tag{3.9}$$

$$\mu_2 = \frac{m_{22}-m_{11}-\sqrt{(m_{22}-m_{11})^2+4m_{12}m_{21}}}{2m_{12}}, \tag{3.10}$$

in the original coordinates, for $m_{12} \neq 0$. These lines divide $\mathbb{R}^2$ into regions where the form is of distinct sign. The vector $\tilde{a}$ must lie in a region of non-negative sign and $\tilde{b}$ must lie in a region of non-positive sign, with the counterclockwise angle from $\tilde{a}$ to $\tilde{b}$ being at most 180° and determining the cone K.

Example 14: Consider the simple example reaction mechanism

$$\alpha_1A_1+\alpha_2A_2 \xrightarrow{k} \bar{\alpha}_1A_1+\bar{\alpha}_2A_2. \tag{3.11}$$

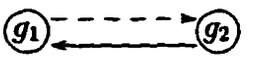
From page 27 of [18], we know that (3.11) does not induce an order preserving flow if $\bar{\alpha}_1 > \alpha_1$ and $\alpha_2 > \bar{\alpha}_2$. Let $A_1(t)$ and $A_2(t)$ denote the concentrations of species A_1 and A_2 at time t. Reaction (3.11) induces the system of differential equations

$$\dot{A}_1(t) = f_1(A_1,A_2) = k(\bar{\alpha}_1-\alpha_1)(A_1(t))^{\alpha_1}(A_2(t))^{\alpha_2}, \text{ and} \tag{3.12}$$

$$\dot{A}_2(t) = f_2(A_1,A_2) = k(\bar{\alpha}_2-\alpha_2)(A_1(t))^{\alpha_1}(A_2(t))^{\alpha_2}, \tag{3.13}$$

subject to initial conditions $A_1(0) = A_{10}$ and $A_2(0) = A_{20}$, where A_{10} and A_{20} are positive for the positivity assumption to hold. The 2×2 Jacobian matrix for

this system has $f_{1,2} > 0$ and $f_{2,1} < 0$, $t \geq 0$. Applying the theory of the previous section, we draw the graph



which consists of one negative directed cycle. The theory of the previous chapter gives us no conclusion. We will try something new.

The system admits the conservation equation

$$A_1(t) - A_{10} = \frac{\bar{\alpha}_1 - \alpha_1}{\bar{\alpha}_2 - \alpha_2}(A_2(t) - A_{20}). \tag{3.14}$$

This gives bounds on the concentrations, namely

$$0 \leq A_1(t) \leq A_{10} - \frac{\bar{\alpha}_1 - \alpha_1}{\bar{\alpha}_2 - \alpha_2}A_{20} \text{ and} \tag{3.15}$$

$$0 \leq A_2(t) \leq A_{20} - \frac{\bar{\alpha}_2 - \alpha_2}{\bar{\alpha}_1 - \alpha_1}A_{10}, \tag{3.16}$$

and so our solutions lie in a rectangle. We could conceivably use (3.14) to eliminate $A_2(t)$ from the right hand side of (3.12); this would still leave a complicated differential equation for $A_1(t)$. As can be seen in [18], even seemingly simple monotonicity problems like this can require rather complicated specific arguments and usually require a fair bit of insight into the physical problem. Let us try to apply Theorem 23 instead. From chemical kinetics, $\bar{\alpha}_1 > \alpha_1 > 0$ and $\alpha_2 > \bar{\alpha}_2 \geq 0$ imply that

$$\lim_{t\to\infty} A_2(t) = 0 \text{ and } \lim_{t\to\infty} A_1(t) = A_{10} - \left(\frac{\bar{\alpha}_1 - \alpha_1}{\bar{\alpha}_2 - \alpha_2}\right) A_{20}. \tag{3.17}$$

The corresponding Jacobian matrix is

$$k(A_1(t))^{\alpha_1 - 1}(A_2(t))^{\alpha_2 - 1} \begin{bmatrix} \alpha_1(\bar{\alpha}_1 - \alpha_1)A_2(t) & \alpha_2(\bar{\alpha}_1 - \alpha_1)A_1(t) \\ \alpha_1(\bar{\alpha}_2 - \alpha_2)A_2(t) & \alpha_2(\bar{\alpha}_2 - \alpha_2)A_1(t) \end{bmatrix}. \tag{3.18}$$

Condition (3.7) requires that

$$[\alpha_1(\bar{\alpha}_1 - \alpha_1)A_2(t) + \alpha_2(\bar{\alpha}_2 - \alpha_2)A_1(t)]^2 \geq 0. \tag{3.19}$$

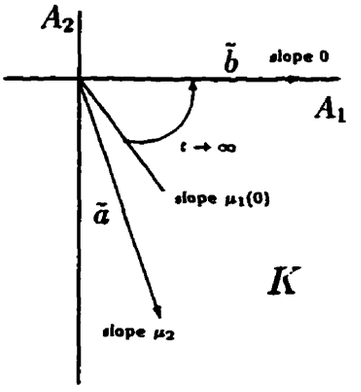


Figure 19: The cone preserved by the Jacobian matrix for Example 14.

Using (3.9)–(3.10), the lines where the corresponding quadratic form vanishes have slope

$$\mu_1 = -\frac{\alpha_1 A_2(t)}{\alpha_2 A_1(t)} \text{ and } \mu_2 = \frac{\bar{\alpha}_2 - \alpha_2}{\bar{\alpha}_1 - \alpha_1}.$$

Using (3.14) and (3.17), we see that

$$\mu_1(0) = -\frac{\alpha_1 A_{20}}{\alpha_2 A_{10}} \leq \mu_1(t) \leq 0, \ \forall\, t \geq 0.$$

Choose

$$\tilde{a} = \begin{bmatrix} 1 \\ \mu_2 \end{bmatrix} \text{ and } \tilde{b} = \tilde{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The quadratic form vanishes along $\tilde{a}$ and is non-positive along $\tilde{b}$, $\forall\, t \geq 0$; hence, the cone K with extreme rays $\tilde{a}$ and $\tilde{b}$ will be preserved by the Jacobian matrix $\forall\, t \geq 0$. Figure 19 shows the cone K; note that $-K$ is also preserved and will lead to the same conclusions. With the cone K so defined, we can apply Theorem 23 with $\tilde{v}(t) = \tilde{A}(t, \tilde{A}_0)$ and $\tilde{w}(t) = \tilde{A}(t, \tilde{A}_0 + \tilde{\beta})$, with $\tilde{\beta} \in K$, and $\tilde{A}(t, \tilde{A}_0)$ the solution to (3.12)–(3.13) with initial condition $\tilde{A}(0) = \tilde{A}_0$. Then $\tilde{w}(t) \geq_K \tilde{v}(t)$, $\forall\, t \geq 0$. In particular, choosing $\tilde{\beta} = \tilde{e}_1$ allows us to conclude that

$$A_1(t, \tilde{A}_0 + \tilde{e}_1) - A_1(t, \tilde{A}_0) \geq 0 \ \Rightarrow \ \frac{\partial A_1}{\partial A_{10}}(t) \geq 0, \ \forall\, t \geq 0, \text{ and} \qquad (3.20)$$

$$A_2(t, \tilde{A}_0 + \tilde{e}_1) - A_2(t, \tilde{A}_0) \leq 0 \;\Rightarrow\; \frac{\partial A_2}{\partial A_{10}}(t) \leq 0, \; \forall\, t \geq 0. \tag{3.21}$$

These same partial derivative results, with strict inequalities, can in fact be obtained through a direct argument.

A variation of Theorem 23 which allows us to use expanding cones will prove useful in the examples. The proof is very similar to the proof presented for Theorem 23, but it is instructive to present it in full. The reader should realize why expanding cones work and why shrinking cones do not work.

Definition 6 *A cone $K(t)$ with extreme rays that change with t is called* **expanding** *if $K(t_1) \subseteq K(t_2)$ whenever $t_1 \leq t_2$.*

Theorem 25 *Consider a closed, convex, solid, expanding cone $K(t)$ in $\mathbb{R}^n$. Suppose $\dot{\tilde{v}} \leq_{K(t)} \tilde{f}(\tilde{v})$, $\dot{\tilde{w}} \geq_{K(t)} \tilde{f}(\tilde{w})$, and $\tilde{w}(0) \geq_{K(0)} \tilde{v}(0)$. Suppose also that $\tilde{f}(\tilde{x})$ is continuously differentiable in $\tilde{x}$ on compact subsets of Ω, Ω open and convex, and that for any compact set, N, $\exists\; l = l(N)$ such that $D\tilde{f}(\tilde{x}) + lI : K(t) \mapsto K(t)$, $\forall\, \tilde{x} \in N$. Then $\tilde{w}(t) \geq_{K(t)} \tilde{v}(t)$, $\forall\, t \geq 0$.*

Proof: Following the proof of Theorem 23, we first suppose that $\dot{\tilde{v}} \leq_{K(t)} \tilde{f}(\tilde{v})$, $\dot{\tilde{w}} >_{K(t)} \tilde{f}(\tilde{w})$, and $\tilde{w}(0) >_{K(0)} \tilde{v}(0)$. We will prove that $\tilde{w}(t) >_{K(t)} \tilde{v}(t)$, $\forall\, t \geq 0$. Suppose this is not the case. Then $\exists$ a first time $t_0 > 0$ such that $\tilde{w}(t) >_{K(t)} \tilde{v}(t)$ for $0 \leq t < t_0$ and $\tilde{w}(t_0) - \tilde{v}(t_0) \in bdy(K(t_0))$. Let N be a compact, convex set containing both trajectories $\tilde{w}(t)$ and $\tilde{v}(t)$ for $0 \leq t < t_0$ (choose a large enough closed ball containing both). Let $\tilde{z}(t) = \tilde{w}(t) - \tilde{v}(t)$. If $\dot{\tilde{z}}(t_0) \in int(K(t_0))$ then $\exists\; t_1$, $0 \leq t_1 < t_0$, such that $\dot{\tilde{z}}(t) \in int(K(t_0))$ for $t_1 < t \leq t_0$ (we are using the fact that $K(t)$ is expanding). Then $\tilde{z}(t) \in int(K(t_0))$ for $0 < t < t_0$ and $\dot{\tilde{z}}(t) \in int(K(t_0))$ for $t_1 < t \leq t_0$ give $\tilde{z}(t_0) \in int(K(t_0))$, a contradiction. Hence, $\dot{\tilde{z}}(t_0) \notin int(K(t_0))$.

By Lemma 20', $\exists$ a supporting hyperplane to $K(t_0)$ at $\tilde{z}(t_0)$. Let $\tilde{b}$ be the outward unit normal to $K(t_0)$ at $\tilde{z}(t_0)$. Since $\tilde{z}(t_0) \in bdy(K(t_0))$ and $\dot{\tilde{z}}(t_0) \notin int(K(t_0))$ we have $\tilde{z}(t_0) \cdot \tilde{b} = 0$ and $\dot{\tilde{z}}(t_0) \cdot \tilde{b} \geq 0$. Now, by the hypotheses, $\dot{\tilde{z}}(t) = \dot{\tilde{w}}(t) - \dot{\tilde{v}}(t) >_{K(t)} \tilde{f}(\tilde{w}) - \tilde{f}(\tilde{v})$, $\forall\ t \geq 0$. Evaluating at t_0 gives $\dot{\tilde{z}}(t_0) - \tilde{f}(\tilde{w}(t_0)) + \tilde{f}(\tilde{v}(t_0)) >_{K(t_0)} \tilde{0}$. Thus,

$$\left[\dot{\tilde{z}}(t_0) - \tilde{f}(\tilde{w}(t_0)) + \tilde{f}(\tilde{v}(t_0))\right] \cdot \tilde{b} < 0 \quad \Rightarrow \quad \left[\tilde{f}(\tilde{w}(t_0)) - \tilde{f}(\tilde{v}(t_0))\right] \cdot \tilde{b} > \dot{\tilde{z}}(t_0) \cdot \tilde{b} \geq 0$$
$$\Rightarrow \quad \left[\tilde{f}(\tilde{w}(t_0)) - \tilde{f}(\tilde{v}(t_0))\right] \notin int(K(t_0))$$

Now,

$$\tilde{f}(\tilde{w}(t_0)) - \tilde{f}(\tilde{v}(t_0)) = \left[\int_0^1 D\tilde{f}(s\tilde{w}(t_0) + (1-s)\tilde{v}(t_0))ds\right] \tilde{z}(t_0).$$

Since, by assumption,

$$\left[\int_0^1 D\tilde{f}(s\tilde{w}(t_0) + (1-s)\tilde{v}(t_0))ds + lI\right] \tilde{z}(t_0) \in K(t_0)$$

we have

$$\left[\int_0^1 D\tilde{f}(s\tilde{w}(t_0) + (1-s)\tilde{v}(t_0))ds + lI\right] \tilde{z}(t_0) \cdot \tilde{b} \leq 0.$$

But

$$\begin{aligned}
&\left[\int_0^1 D\tilde{f}(s\tilde{w}(t_0) + (1-s)\tilde{v}(t_0))ds + lI\right] \tilde{z}(t_0) \cdot \tilde{b} \\
= &\left[\int_0^1 D\tilde{f}(s\tilde{w}(t_0) + (1-s)\tilde{v}(t_0))ds\right] \tilde{z}(t_0) \cdot \tilde{b} \qquad \text{since } \tilde{z}(t_0) \cdot \tilde{b} = 0 \\
= &\left[\tilde{f}(\tilde{w}(t_0)) - \tilde{f}(\tilde{v}(t_0))\right] \cdot \tilde{b} \\
> &\ 0,
\end{aligned}$$

gives a contradiction; hence, $\tilde{w}(t) >_{K(t)} \tilde{v}(t)$, $\forall\ t \geq 0$.

To prove the theorem, again let $\tilde{w}^\epsilon(t) = \tilde{w}(t) + \epsilon e^{\alpha t}\tilde{k}$, where ϵ is a small positive parameter, $\tilde{k}$ is an interior vector of $K(0) \subseteq K(t)$, and α is a constant to be chosen later. Then

$$\tilde{w}^\epsilon(0) = \tilde{w}(0) + \epsilon\tilde{k} >_{K(0)} \tilde{w}(0) \geq_{K(0)} \tilde{v}(0).$$

Now, let X be the compact set $\{\tilde{f}(\tilde{w}(t)) - \tilde{f}(\tilde{w}^\epsilon(t)) : t \in [0,T]\}$. Lemma 21 tells us that for some $\beta > 0$ and for all $t \in [0,T]$

$$\tilde{f}(\tilde{w}(t)) - \tilde{f}(\tilde{w}^\epsilon(t)) + \beta\tilde{k} >_{K(0)} \tilde{0} \;\Rightarrow\; \tilde{f}(\tilde{w}(t)) - \tilde{f}(\tilde{w}^\epsilon(t)) + \beta\tilde{k} >_{K(t)} \tilde{0},$$

since $K(t)$ is expanding. Now, choosing $\alpha = \frac{\beta}{\epsilon}$, we have for $t \in [0,T]$

$$\dot{\tilde{w}}^\epsilon(t) = \dot{\tilde{w}}(t) + \epsilon\alpha e^{\alpha t}\tilde{k} \geq_{K(t)} \tilde{f}(\tilde{w}(t)) + \beta\tilde{k} >_{K(t)} \tilde{f}(\tilde{w}^\epsilon(t)).$$

By the first result in this proof, we can conclude that $\tilde{v}(t) <_{K(t)} \tilde{w}^\epsilon(t) = \tilde{w}(t) + \epsilon e^{\alpha t}\tilde{k}$, for $t \in [0,T]$. Letting $\epsilon \to 0$ proves the theorem. □

Remark: Theorem 25 could be stated for non-solid cones, where we would additionally demand that the smallest subspace containing $K(t)$ be the same for each $t \geq 0$. In our example applications, we will only need the result for solid cones.

To illustrate that shrinking cones which satisfy the hypotheses of Theorem 25 will not lead us to the conclusion of the theorem, consider the following example.

Example 15: Suppose that our system, $\dot{\tilde{x}} = \tilde{f}(\tilde{x})$, is given by

$$\begin{aligned} \dot{x}_1 &= x_1 \Rightarrow x_1(t) = x_1(0)e^t \\ \dot{x}_2 &= x_2 \Rightarrow x_2(t) = x_2(0)e^t, \end{aligned}$$

with easily calculable solutions. In this case $D\tilde{f}$ is just the identity matrix. Consider the cone $K(t)$ with extreme rays $\tilde{a} = [1,0]^T$ and $\tilde{b}(t) = [1,e^{-t}]^T$, as illustrated in Figure 20. Since $(D\tilde{f}+lI)\tilde{a} = (1+l)\tilde{a} \in K(t)$ and $(D\tilde{f}+lI)\tilde{b}(t) = (1+l)\tilde{b}(t) \in K(t)$, the essential hypothesis is satisfied. Now, let $\tilde{w}(t)$ and $\tilde{v}(t)$ be two solutions with initial values

$$\tilde{w}(0) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \geq_{K(0)} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \tilde{v}(0);$$

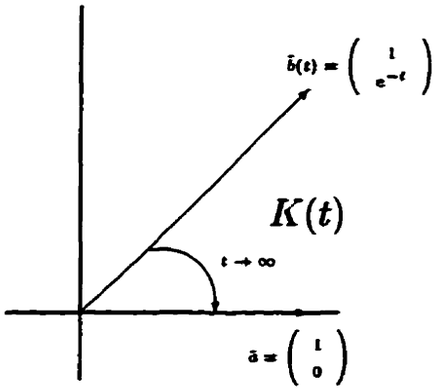


Figure 20: A shrinking cone $K(t)$ with extreme rays $\tilde{a}$ and $\tilde{b}(t)$.

then the solutions satisfy

$$\tilde{w}(t) - \tilde{v}(t) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} e^t - \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^t = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^t \notin K(t) \text{ for } t > 0.$$

It is perhaps interesting to note that this example suffers from a richness of possibilities: $D\tilde{f}$ preserves any cone we choose! We could use the machinery of the previous chapter, use the machinery of this chapter with the positive quadrant as our cone, or just take partial derivatives of our solutions to conclude that the matrix of sign patterns for the partial derivatives with respect to initial conditions is given in Table 7, where the entries have their usual meanings. This should highlight the difficulty in finding cones that yield valuable results.

	$x_1(t)$	$x_2(t)$
$x_{1,0}$	++	0
$x_{2,0}$	0	++

Table 7: Signs of partial derivatives with respect to initial concentrations for Example 15.

In the previous chapter on monotonicity with respect to an orthant, we saw that graph theory played a key role in establishing strict sign results. When dealing with cones, it is perhaps not all together obvious how to use graphs to this end or devise strict sign results some other way. We present the following theorem as a first step; it will lead us into a graph theoretic discussion. This theorem could also be stated for expanding cones.

Theorem 26 *Let K be a closed, convex cone in $\mathbf{R}^n$. Suppose that $\dot{\tilde{v}} \leq_K \tilde{f}(\tilde{v})$, $\dot{\tilde{w}} \geq_K \tilde{f}(\tilde{w})$, and $\tilde{w}(0) \geq_K \tilde{v}(0)$. Suppose also that $\tilde{f}(\tilde{x})$ is continuously differentiable in $\tilde{x}$ on compact subsets of $\mathbf{R}^n$ and that for any compact set, N, $\exists\ l = l(N)$ such that*

$$D\tilde{f}(\tilde{x}) + lI : K \mapsto K,\ \forall\ \tilde{x} \in N.$$

Further assume that for N and l chosen as above, $\exists$ a positive integer m such that

$$\left(D\tilde{f}(\tilde{x}) + lI\right)^m : K \setminus \{0\} \mapsto relint(K),\ \forall\ \tilde{x} \in N.$$

Then $\tilde{w}(t) >_K \tilde{v}(t)$, $\forall\ t > 0$.

Proof: Let N be a compact set containing both trajectories $\tilde{w}(t)$ and Let N be a compact, convex set containing both trajectories $\tilde{w}(t)$ and $\tilde{v}(t)$ for $0 \leq t < T$ (choose a large enough closed ball containing both), and let $\tilde{z}(t) = \tilde{w}(t) - \tilde{v}(t)$; then $\tilde{z}(0) \geq_K \tilde{0}$ and, by Lemma 22, $\tilde{z}(t) \in \pi$ for $t > 0$, where π is the smallest subspace of $\mathbf{R}^n$ containing K. Furthermore,

$$\begin{aligned}
\dot{\tilde{z}} &\geq_K \tilde{f}(\tilde{w}) - \tilde{f}(\tilde{v}) \\
&= \tilde{f}(\tilde{z} + \tilde{v}) - \tilde{f}(\tilde{v}) \\
&= \left[\int_0^1 D\tilde{f}(s\tilde{z}(t) + \tilde{v}(t))ds\right]\tilde{z}(t).
\end{aligned}$$

Choose l as in the hypothesis. We have

$$\dot{\tilde{z}} + l\tilde{z} \geq_K \left[\int_0^1 D\tilde{f}(s\tilde{z}(t) + \tilde{v}(t))ds + lI\right] \tilde{z}(t)$$

$$\Rightarrow \frac{d}{dt}\tilde{z}e^{lt} \geq_K e^{lt}\left[\int_0^1 D\tilde{f}(s\tilde{z}(t) + \tilde{v}(t))ds + lI\right] \tilde{z}(t).$$

Integrate with respect to t from t_0 to t, $0 \leq t \leq T$, to get

$$\tilde{z} \geq_K e^{l(t_0-t)}\tilde{z}(t_0) + \int_{t_0}^t e^{l(s_2-t)}\left[\int_0^1 D\tilde{f}(s_1\tilde{z}(s_2) + \tilde{v}(s_2))ds_1 + lI\right] \tilde{z}(s_2)ds_2. \quad (3.22)$$

Note that we are using the property that if $\tilde{q}(t) \geq_K \tilde{r}(t)$, for $0 \leq t \leq T$, then

$$\int_{t_0}^t \tilde{q}(s)ds \geq_K \int_{t_0}^t \tilde{r}(s)ds,$$

for $0 \leq t_0 \leq t < T$, proven with Riemann integral and cone closure under addition. So,

$$\tilde{z} \geq_K e^{-l(t-t_0)}\tilde{z}(t_0) + \int_{t_0}^t e^{-l(t-s_2)}M(s_2)\tilde{z}(s_2)ds_2, \quad (3.23)$$

$t_0 \leq t < T$, where

$$M(s_2) = \int_0^1 \left[D\tilde{f}(s_1\tilde{z}(s_2) + \tilde{v}(s_2)) + lI\right] ds_1.$$

By the Extended Kamke-Müller Theorem (Theorem 23), $\tilde{z}(t) \in K$, for $0 \leq t < T$. Furthermore, if $\tilde{z}(t_0) \in \mathit{relint}(K)$ then (3.23) tells us that $\tilde{z}(t) \in \mathit{relint}(K)$, $t_0 \leq t < T$, since $M(s_2)\tilde{z}(s_2) \in K$ because $D\tilde{f}$ preserves K.

We plan to apply inequality (3.23) to the $z(s_2)$ term on the right hand side of (3.23). We have (3.23) with $t_0 = 0$:

$$\begin{aligned}\tilde{z}(t) &\geq_K e^{-lt}\tilde{z}(0) + \int_0^t e^{-l(t-s_2)}M(s_2)\tilde{z}(s_2)ds_2, \\ &\geq_K \int_0^t e^{-l(t-s_2)}M(s_2)\tilde{z}(s_2)ds_2, \text{ since } \tilde{z}(0) \in K. \quad (3.24)\end{aligned}$$

This gives

$$\tilde{z}(s_2) \geq_K \int_0^{s_2} e^{-l(s_2-s_3)} M(s_3)\tilde{z}(s_3)ds_3,\ 0 \leq s_2 \leq t,$$

or, in other words,

$$\tilde{z}(s_2) - \int_0^{s_2} e^{-l(s_2-s_3)} M(s_3)\tilde{z}(s_3)ds_3 \in K,$$

which, in turn, gives

$$M(s_2)\left[z(\tilde{s}_2) - \int_0^{s_2} e^{-l(s_2-s_3)} M(s_3)\tilde{z}(s_3)ds_3\right] \in K,$$

since $M(s_2)$ preserves K. So,

$$\begin{aligned} M(s_2)\tilde{z}(s_2) &\geq_K M(s_2)\int_0^{s_2} e^{-l(s_2-s_3)} M(s_3)\tilde{z}(s_3)ds_3 \\ &= \int_0^{s_2} e^{-l(s_2-s_3)} M(s_2)M(s_3)\tilde{z}(s_3)ds_3, \end{aligned}$$

giving

$$\int_0^t e^{-(t-s_2)} M(s_2)\tilde{z}(s_2)ds_2 \geq_K \int_0^t \int_0^{s_2} e^{-l(t-s_3)} M(s_2)M(s_3)\tilde{z}(s_3)ds_3ds_2,$$

or

$$\tilde{z}(t) \geq_K \int_0^t \int_0^{s_2} e^{-l(t-s_3)} M(s_2)M(s_3)\tilde{z}(s_3)ds_3ds_2.$$

We can repeatedly iterate in this way to get

$$\tilde{z}(t) \geq_K \int_0^t \int_0^{s_2} \cdots \int_0^{s_n} e^{-l(t-s_n)} M(s_2)\cdots M(s_{n+1})\tilde{z}(s_{n+1})ds_{n+1}\cdots ds_2, \quad (3.25)$$

$0 \leq s_{n+1} \leq \cdots \leq s_2 \leq t$. We consider the m^{th} iterate, m chosen as in the final hypothesis of the theorem. For t sufficiently small, we get

$$\begin{aligned} \tilde{z}(t) &\geq_K e^{-lt}\int_0^t \int_0^{s_2} \cdots \int_0^{s_m} M(0)\cdots M(0)\tilde{z}(0)(1+o(1))ds_{m+1}\cdots ds_2 \\ &= e^{-lt}\left[(M(0))^m \tilde{z}(0)\frac{t^m}{m!}(1+o(1))\right] \\ &>_K \tilde{0}, \end{aligned}$$

using the final hypothesis. This means that there is a small t_0, $0 < t_0 < T$, such that $\tilde{z}(t_0) \in relint(K)$. As stated earlier, (3.23) then implies $\tilde{z}(t) \in relint(K)$, $t_0 \leq t \leq T$. Since t_0 can be chosen as small as we like, we conclude that $\tilde{z}(t) \in relint(K)$, $0 < t \leq T$.

□

Remark: As with earlier results, Theorem 26 is stated for a closed, convex cone K. The additional hypothesis for strong monotonicity, namely that

$$(D\tilde{f}(\tilde{x}) + lI)^m : K \setminus \{\tilde{0}\} \mapsto relint(K),$$

will not hold for an unpointed cone. An unpointed cone will contain a subspace of $\mathbb{R}^n$; so, there will be two vectors $\tilde{e}$ and $-\tilde{e}$ on the boundary of the cone. It is clear that $(D\tilde{f}(\tilde{x}) + lI)^m$ cannot map $\tilde{e}$ to $relint(K)$, since it would then map $-\tilde{e}$ to the exterior of K. In fact, the only way for $D\tilde{f} + lI$ to map such a cone into itself is for the subspace to be invariant under the transformation.

From the proof of Theorem 26, we get the following corollary.

Corollary 27 *Under the hypotheses of Theorem 26,*

$$\tilde{w}(t_0) >_K \tilde{v}(t_0) \Rightarrow \tilde{w}(t) >_K \tilde{v}(t),\ t \geq t_0.$$

Writing

$$\frac{\partial \tilde{x}}{\partial \tilde{k}} = D_{\tilde{x}_0}\tilde{x} \cdot \tilde{k},\ \frac{\partial \tilde{x}}{\partial \tilde{k}}(0) = \tilde{k}, \tag{3.26}$$

leads to the following corollary to Theorem 26.

Corollary 28 *Suppose that $\tilde{f}(\tilde{x})$ is continuously differentiable in $\tilde{x}$ on compact subsets of $\mathbb{R}^n$ and that for any compact set, N, $\exists\ l = l(N)$ such that*

$$D\tilde{f}(\tilde{x}) + lI : K \mapsto K,\ \forall\ \tilde{x} \in N,$$

where K is a closed, convex cone in $\mathbb{R}^n$. Further assume that for N and l chosen as above, $\exists$ a positive integer m such that

$$\left(D\tilde{f}(\tilde{x}) + lI\right)^m \tilde{k} \in relint(K),\ \forall\ \tilde{x} \in N,$$

for some vector $\tilde{k} \in K \setminus \{\tilde{0}\}$; then

$$\frac{\partial \tilde{x}}{\partial \tilde{k}} >_K \tilde{0},$$

for $t > 0$ ($t \geq 0$ if $\tilde{k} >_K \tilde{0}$).

Proof: Differentiate equation (2.7) with respect to $\tilde{x}_0$ to get

$$\frac{d}{dt}(D_{\tilde{x}_0}\tilde{x}) = D_{\tilde{x}_0}\tilde{f}(\tilde{x})D_{\tilde{x}_0}\tilde{x}, \tag{3.27}$$

which, upon dotting with $\tilde{k}$, gives

$$\frac{d}{dt}\left(\frac{\partial \tilde{x}}{\partial \tilde{k}}\right) = D_{\tilde{x}_0}\tilde{f}(\tilde{x})\frac{\partial \tilde{x}}{\partial \tilde{k}}.$$

For convenience, let $\tilde{w}(t) = \frac{\partial \tilde{x}}{\partial \tilde{k}}$; then $\dot{\tilde{w}} = D\tilde{f}(\tilde{x})\tilde{w}$ and $\tilde{w}(0) = \tilde{k}$. With l chosen as in the corollary, we write

$$\dot{\tilde{w}} + l\tilde{w} = (D\tilde{f}(\tilde{x}) + lI)\tilde{w},$$

which we solve to get

$$\tilde{w}(t) = e^{-lt}\tilde{w}(t_0) + \int_{t_0}^{t} e^{-l(t-s)}(D\tilde{f}(\tilde{x}(s)) + lI)\tilde{w}(s)ds. \tag{3.28}$$

We notice that $\tilde{w}(t_0) >_K \tilde{0}$ implies that $\tilde{w}(t) >_K \tilde{0}$, $t \geq t_0$, since each term on the right hand side is in K. As in the theorem, with $t_0 = 0$, we now iterate, repeatedly replacing the $\tilde{w}$ on the right hand side by the entire expression, which at iteration m, chosen as in the corollary, gives us

$$\tilde{w}(t) = e^{-lt}\tilde{w}(0) + \int_0^t \int_0^{s_2} \cdots \int_0^{s_n} e^{-l(t-s_n)}M(s_2)\cdots M(s_{n+1})\tilde{w}(s_{n+1})ds_{n+1}\cdots ds_2,$$

where $M(s) = (D\tilde{f}(\tilde{x}(s)) + lI)$ and $0 \leq s_{n+1} \leq \cdots \leq s_2 \leq t$. For t sufficiently small, we get

$$\begin{aligned}
\tilde{w}(t) &= e^{-lt}\tilde{w}(0) + \int_0^t \int_0^{s_2} \cdots \int_0^{s_n} e^{-lt} M(0) \cdots M(0)\tilde{w}(0)(1 + o(1))ds_{n+1} \cdots ds_2 \\
&= e^{-lt}\tilde{w}(0) + e^{-lt}\left[(M(0))^m \tilde{w}(0)\frac{t^m}{m!}(1 + o(1))\right] \\
&= e^{-lt}\tilde{k} + e^{-lt}\left[(M(0))^m \tilde{k}\frac{t^m}{m!}(1 + o(1))\right] \\
&>_K \tilde{0}, \qquad (3.29)
\end{aligned}$$

since $(M(0))^m\tilde{k} >_K \tilde{0}$, using the second hypothesis. Applying Corollary 27 lead to $\tilde{w}(t) >_K \tilde{0}, t > 0$. □

In applications, $D\tilde{f}(\tilde{x}) + lI$ may change substantially from time zero to time greater than zero. For example, solutions to chemical kinetics or epidemiological problems may begin with some components zero, but, if the positivity assumption holds, all components will be positive for positive time. Under these circumstances, it is possible that the strong monotonicity hypothesis of Corollary 28 will not be satisfied at $t = 0$. We formulate another corollary that deals specifically with the applications we plan to analyze.

Corollary 29 *Assume that $\bar{\mathcal{O}}$ is positively invariant and that the positivity assumption holds ($\tilde{x}(t) > \tilde{0}$ for $t > 0$). Suppose that $\tilde{f}(\tilde{x})$ is continuously differentiable on $\bar{\mathcal{O}}$ and that $\exists\ l$ such that*

$$D\tilde{f}(\tilde{x}) + lI : K \mapsto K,\ \forall\ \tilde{x} \in \bar{\mathcal{O}},$$

where K is a closed, convex cone in $\mathbf{R}^n$. Further assume that for l chosen as above, $\exists$ a positive integer m such that

$$\left(D\tilde{f}(\tilde{x}(t_0)) + lI\right)^m \tilde{k} \in relint(K),$$

for some vector $\tilde{k} \in K \setminus \{\tilde{0}\}$*; then*

$$\frac{\partial \tilde{x}}{\partial \tilde{k}} >_K \tilde{0},$$

for $t > t_0$ ($t \geq t_0$ *if* $\tilde{k} >_K \tilde{0}$).

Proof: From the proof of Corollary 28, we have that

$$\tilde{w}(t) = e^{-lt}\tilde{k} + \tilde{g}(t),$$

where $\tilde{g}(t) \in K$.

We follow the same line of attack as in the proof of Corollary 28, iterating to get

$$\tilde{w}(t) = e^{-lt}\tilde{w}(t_0) + \int_{t_0}^{t}\int_{t_0}^{s_2}\cdots\int_{t_0}^{s_n} e^{-l(t-s_n)}M(s_2)\cdots M(s_{n+1})\tilde{w}(s_{n+1})ds_{n+1}\cdots ds_2,$$

where $M(s) = (D\tilde{f}(\tilde{x}(s)) + lI)$ and $t_0 \leq s_{n+1} \leq \cdots \leq s_2 \leq t$. For t sufficiently close to t_0, we get

$$\begin{aligned}
\tilde{w}(t) &= e^{-lt}\tilde{w}(t_0) \\
&\quad + \int_{t_0}^{t}\int_{t_0}^{s_2}\cdots\int_{t_0}^{s_n} e^{-lt}M(t_0)\cdots M(t_0)\tilde{w}(t_0)(1+o(1))ds_{n+1}\cdots ds_2 \\
&= e^{-lt}\tilde{w}(t_0) + e^{-lt}\left[(M(t_0))^m\tilde{w}(t_0)\frac{(t-t_0)^m}{m!}(1+o(1))\right] \\
&= e^{-lt}\left(e^{-lt_0}\tilde{k} + \tilde{g}(t_0)\right) \\
&\quad + e^{-lt}\left((M(t_0))^m(e^{-lt_0}\tilde{k} + \tilde{g}(t_0))\frac{(t-t_0)^m}{m!}(1+o(1))\right) \\
&>_K \tilde{0}, \qquad (3.30)
\end{aligned}$$

since $(M(t_0))^m\tilde{k} >_K \tilde{0}$, by the second hypothesis. Apply Corollary 27 to conclude $\tilde{w}(t) >_K \tilde{0}$, $t > t_0$ ($t \geq t_0$ if $\tilde{k} >_K \tilde{0}$). □

Before using the ideas of Theorem 26 to motivate a graph theoretic approach to monotonicity with respect to general cones, we first discuss the place of condition (3.3) amidst similar conditions in the literature.

3.2.1 A Discussion of the Cone Preserving Condition

Having established several results involving condition (3.3), we now present several related conditions that are present in the literature:

(V1) $\tilde{v} \leq_K \tilde{w}$ and $\psi(\tilde{v}) = \psi(\tilde{w}) \Rightarrow \psi(\tilde{f}(\tilde{v})) \leq \psi(\tilde{f}(\tilde{w}))$,

(V2) $\tilde{w} - \tilde{v} \in relbdy(K)$ and $\psi(\tilde{v}) = \psi(\tilde{w}) \Rightarrow \psi(\tilde{f}(\tilde{v})) \leq \psi(\tilde{f}(\tilde{w}))$,

(V2′) $\forall\ \tilde{x} \in N$, N compact, $(D\tilde{f}(\tilde{x}))\tilde{e} \cdot \tilde{N}_{\tilde{e}} \geq 0$, $\forall\ \tilde{e} \in relbdy(K)$, where $\tilde{e} \cdot \tilde{N}_{\tilde{e}} = 0$, $\tilde{N}_{\tilde{e}} \in K^*$,

(K1) $\forall\ \tilde{x} \in N$, N compact, $\exists\ l$ such that $D\tilde{f}(\tilde{x}) + lI : K \mapsto K$, and

(W1) $\forall\ \tilde{x}, \tilde{y} \in N$, $\exists\ \lambda > 0$ such that $\tilde{x} \leq_K \tilde{y} \Rightarrow \tilde{f}(\tilde{x}) + \lambda\tilde{x} \leq_K \tilde{f}(\tilde{y}) + \lambda\tilde{y}$.

(V1) and (V2) are due to Volkmann (see [37]); in each case ψ is a functional on K with $\psi(\tilde{x}) \geq 0$ for $\tilde{x} \in K$. (V2) is also given in [23]. (V2′) is the analog to (V2) for differentiable $\tilde{f}$; it is not stated in Volkmann, but it follows naturally from (V2). (W1) is due to Walter (see [40] or [29]). The research of Walter and Volkmann is abstract, dealing with functionals in Banach spaces.

To connect these conditions, we state the following theorem.

Theorem 30 *For $\tilde{f}(\tilde{x})$ continuously differentiable and K a closed, convex cone,*

$$(V1) \Longleftrightarrow (V2) \Longleftrightarrow (V2') \Longleftarrow (K1) \Longleftrightarrow (W1).$$

If K is polyhedral, then

$$(V2') \Longrightarrow (K1).$$

Proof: (K1) $\Rightarrow$ (V2′): Suppose that (K1) holds; then $\forall\ \tilde{x}$ in some compact set N, $\exists\ l$ such that $D\tilde{f}(\tilde{x}) + lI : K \mapsto K$. For $\tilde{e} \in relbdy(K)$, $(D\tilde{f}(\tilde{x}) + lI)\tilde{e} \cdot \tilde{N}_{\tilde{e}} \geq 0$, implying $(D\tilde{f}(\tilde{x}))\tilde{e} \cdot \tilde{N}_{\tilde{e}} \geq 0$ and proving (V2′).

(V2′) $\Rightarrow$ (V2): Suppose that (V2′) holds. Choose $\tilde{w}$ and $\tilde{v}$ such that let $\tilde{e} = \tilde{w} - \tilde{v} \in$ *relbdy*($\tilde{K}$); then $\forall\ \tilde{x}$ in some compact set N (choose N large enough to contain the line segment connecting $\tilde{w}$ and $\tilde{v}$),

$$(D\tilde{f}(\tilde{x}))(\tilde{w} - \tilde{v}) \cdot \tilde{N}_{\tilde{e}} \geq 0. \tag{3.31}$$

Define $\psi(\tilde{z}) = \tilde{N}_{\tilde{e}} \cdot \tilde{z}$, so $\psi(\tilde{w} - \tilde{v}) = 0$, implying $\psi(\tilde{v}) = \psi(\tilde{w})$. For (V2) to hold, we must have

$$(\tilde{f}(\tilde{w}) - \tilde{f}(\tilde{v})) \cdot \tilde{N}_{\tilde{e}} \geq 0,$$

or, using the mean value theorem,

$$\int_0^1 \left[D\tilde{f}(s\tilde{w} + (1-s)\tilde{v})ds\right](\tilde{w} - \tilde{v}) \cdot \tilde{N}_{\tilde{e}} \geq 0,$$

Evaluating (3.31) at $s\tilde{w} + (1-s)\tilde{v}$, $s \in [0,1]$, gives

$$D\tilde{f}(s\tilde{w} + (1-s)\tilde{v})(\tilde{w} - \tilde{v}) \cdot \tilde{N}_{\tilde{e}} \geq 0,$$

implying

$$\int_0^1 \left[D\tilde{f}(s\tilde{w} + (1-s)\tilde{v})ds\right](\tilde{w} - \tilde{v}) \cdot \tilde{N}_{\tilde{e}} \geq 0$$

and proving (V2).

(V2) $\Rightarrow$ (V1): Suppose that (V2) holds; then $\psi(\tilde{w}) = \psi(\tilde{v}) \Rightarrow \exists\ \tilde{b}$ with $\psi(\tilde{w}) = \tilde{b} \cdot \tilde{w}$, so $\tilde{b} \cdot (\tilde{w} - \tilde{v}) = \psi(\tilde{w} - \tilde{v}) = 0$. Since $\tilde{w} - \tilde{v} \in K$, $\tilde{w} - \tilde{v} = \sum_i \lambda_i \tilde{v}_i$, $\tilde{v}_i \in$ *relbdy*($\tilde{K}$), $\sum_i \lambda_i = 1$, $\lambda_i \geq 0$, $\tilde{v}_i$ linearly independent, $\psi(\tilde{v}_i) \geq 0$; but $\psi(\tilde{w} - \tilde{v}) = 0$ implies that $\psi(\tilde{v}_i) = 0$, $\forall\ i$. Putting things together,

$$\begin{aligned}
\tilde{f}(\tilde{w}) &= \tilde{f}\left(\tilde{v} + \sum_{i=1}^{n} \lambda_i \tilde{v}_i\right) \\
&= \tilde{f}\left(\tilde{v} + \lambda_n \tilde{v}_n + \sum_{i=1}^{n-1} \lambda_i \tilde{v}_i\right) \\
&= \tilde{f}(\tilde{w}_1 + \lambda_n \tilde{v}_n),
\end{aligned}$$

where $\tilde{w}_j = \tilde{v} + \sum_{i=1}^{n-j} \lambda_i \tilde{v}_i$; hence,

$$\begin{aligned}
\psi(\tilde{f}(\tilde{w})) &= \psi(\tilde{f}(\tilde{w}_1 + \lambda_n \tilde{v}_n)) \\
&\geq \psi(\tilde{f}(\tilde{w}_1)), \text{ using } (V2), \\
&= \tilde{f}(\tilde{w}_2 + \lambda_{n-1}\tilde{v}_{n-1}), \\
&\geq \psi(\tilde{f}(\tilde{w}_2)), \text{ using } (V2).
\end{aligned}$$

Continuing in this way leads to $\psi(\tilde{f}(\tilde{w})) \geq \psi(\tilde{f}(\tilde{v}))$, proving (V1).

(V1) $\Rightarrow$ (V2): Suppose that (V1) holds; since $\tilde{w} - \tilde{v} \in relbdy(K) \Rightarrow \tilde{v} \leq_K \tilde{w}$, (V2) holds.

(V2) $\Rightarrow$ (V2′): Suppose that (V2) holds; Choose $\tilde{v}$ arbitrarily and let $\tilde{e}$ be any vector in $relbdy(K)$. Define $\tilde{w}$ by $\tilde{w} - \tilde{v} = \epsilon\tilde{e}$, $\epsilon > 0$, and choose $\psi(\tilde{z}) = \tilde{N}_{\tilde{e}} \cdot \tilde{z}$. Then

$$\psi(\tilde{w} - \tilde{v}) = \tilde{N}_{\tilde{e}} \cdot (\tilde{w} - \tilde{v}) = \tilde{N}_{\tilde{e}} \cdot \epsilon\tilde{e} = 0,$$

implying that $\psi(\tilde{w}) = \psi(\tilde{v})$. Since (V2) holds,

$$\begin{aligned}
&\psi(\tilde{f}(\tilde{v})) \leq_K \psi(\tilde{f}(\tilde{w})) = \psi(\tilde{f}(\tilde{v} + \epsilon\tilde{e})) \\
&\Rightarrow \frac{\psi(\tilde{f}(\tilde{v} + \epsilon\tilde{e})) - \psi(\tilde{f}(\tilde{v}))}{\epsilon} \geq_K 0 \\
&\Rightarrow \left[\frac{\tilde{f}(\tilde{v} + \epsilon\tilde{e}) - \tilde{f}(\tilde{v})}{\epsilon}\right] \cdot \tilde{N}_{\tilde{e}} \geq 0 \\
&\Rightarrow (D\tilde{f}(\tilde{v}))\tilde{e} \cdot \tilde{N}_{\tilde{e}} \geq 0,
\end{aligned}$$

proving (V2′).

(W1) $\Rightarrow$ (K1): Suppose that (W1) holds; let $\tilde{y} = \tilde{x} + s\tilde{k}$, $\tilde{k} \in K$, $s > 0$. It follows from (W1) that

$$\tilde{f}(\tilde{y}) - \tilde{f}(\tilde{x}) + \lambda(\tilde{y} - \tilde{x}) \geq_K \tilde{0},$$

which gives

$$\frac{\tilde{f}(\tilde{x}+s\tilde{k})-\tilde{f}(\tilde{x})+\lambda s\tilde{k}}{s} \geq_K \tilde{0}.$$

Letting $s \to 0$, we get

$$(D\tilde{f}(\tilde{x}))\tilde{k}+\lambda\tilde{k} \geq_K \tilde{0},$$

which means that

$$(D\tilde{f}(\tilde{x})+\lambda I) : K \mapsto K,$$

proving (K1).

(K1) $\Rightarrow$ (W1): Suppose that (K1) holds; we follow the previous argument in reverse. Let $\tilde{k} \in K$ and $\tilde{k}^* \in K^*$. (K1) tells us that $\forall\ \tilde{x}$ in some compact set N,

$$\left[(D\tilde{f}(\tilde{x}+s\tilde{k}))\tilde{k}+l\tilde{k}\right]\cdot\tilde{k}^* \geq 0, \tag{3.32}$$

where we have strategically chosen the argument of $D\tilde{f}$ and $s \in [0,1]$. Notice that if $\tilde{g}(s)=\tilde{F}(\tilde{x}+s\tilde{k})$, then $\tilde{g}'(s)=D\tilde{F}(\tilde{x}+s\tilde{k})\tilde{k}$; hence, if we define

$$\tilde{g}(s)=\left[\tilde{f}(\tilde{x}+s\tilde{k})+sl\tilde{k}\right]\cdot\tilde{k}^*,$$

then (3.32) says that $\tilde{g}'(s) \geq 0$. We can conclude that $\tilde{g}(1) \geq \tilde{g}(0)$, say. This gives

$$\left[\tilde{f}(\tilde{x}+\tilde{k})+l\tilde{k}-\tilde{f}(\tilde{x})\right]\cdot\tilde{k}^* \geq 0,$$

which, upon letting $\tilde{y}=\tilde{x}+\tilde{k} \geq_K \tilde{x}$, leads to

$$\tilde{f}(\tilde{y})-\tilde{f}(\tilde{x})+l(\tilde{y}-\tilde{x}) \geq_K \tilde{0}.$$

Rearranging proves (W1).

Finally, we prove that if K is polyhedral, then (V2′) $\Rightarrow$ (K1): Suppose (V2′) holds. Since K is polyhedral, we need only consider the finite number of generators of K, denoted $\tilde{e}^i$, $i=1,\ldots,n_k$. For a given $\tilde{e}^i$, there is a set of choices for $\tilde{N}_{\tilde{e}^i} \in K^*$

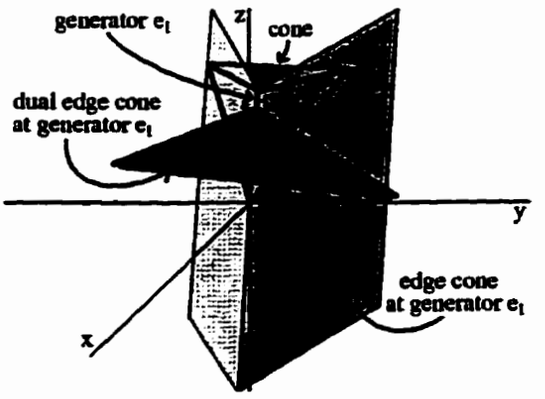


Figure 21: An edge cone and dual edge cone in $\mathbb{R}^3$.

with $\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^i} = 0$. We label this set $N(K, \tilde{e}^i)$ and notice that it is a polyhedral *dual edge cone* of K. As the name implies, $N(K, \tilde{e}^i)$ is dual to the *edge cone* of K based on $\tilde{e}^i$. See Figure 21. Since $N(K, \tilde{e}^i)$ is also polyhedral, we need only consider its finite number of generators, denoted $\tilde{N}_{\tilde{e}^i,j}$, $i = 1, \ldots, n_k$, $j = 1, \ldots, n_{k_j^*}$, since any $\tilde{N}_{\tilde{e}^i} \in K^*$ with $\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^i} = 0$ is generated by the $\tilde{N}_{\tilde{e}^i,j}$'s. This labelling is not unique, but it suits our needs. Since (V2′) holds, for each i $\exists$ compact N, such that $\forall\ \tilde{x} \in N$,

$$(D\tilde{f}(\tilde{x}))\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^i,j} \geq 0,$$

where $\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^i,j} = 0$, $j = 1, \ldots, n_{k_j^*}$.

In order for $D\tilde{f}(\tilde{x}) + lI : K \mapsto K$ we need $(D\tilde{f}(\tilde{x}) + lI)\tilde{e}^i \in K$ for each i. This is true if

$$r(\tilde{x}, i, j, k, l) = (D\tilde{f}(\tilde{x}) + lI)\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^k,j} \geq 0,$$

for $j = 1, \ldots, n_{k_j^*}$ and $\forall\ k$; that is, if the image of $\tilde{e}^i$ has non-negative inner product with all of the generators of K^*. Expanding gives

$$r(\tilde{x}, i, j, k, l) = D\tilde{f}(\tilde{x})\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^k,j} + l\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^k,j}.$$

For $k = i$,

$$r(\tilde{x}, i, j, i, l) \;=\; D\tilde{f}(\tilde{x})\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^i,j} + l\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^i,j}$$

$$\begin{aligned} &= D\tilde{f}(\tilde{x})\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^{i,j}} \\ &\geq 0, \end{aligned}$$

by (V2′). For $k \neq i$, notice that $\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^{k,j}} \geq 0$ because $\tilde{e}^i \in K$ and $\tilde{N}_{\tilde{e}^{k,j}} \in K^*$. We examine the two cases.

(1) $\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^{k,j}} > 0$. Picking

$$l_{i,j,k} = \frac{1}{\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^{k,j}}} \max_{\tilde{x} \in N} \left| D\tilde{f}(\tilde{x})\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^{k,j}} \right|$$

gives $r(\tilde{x}, i, j, i, l) \geq 0$.

(2) $\tilde{e}^i \cdot \tilde{N}_{\tilde{e}^{k,j}} = 0$. This means that $\tilde{N}_{\tilde{e}^{k,j}} = \tilde{N}_{\tilde{e}^{i,j'}}$ for some $j' = 1, \ldots, n_{k_j^*}$; hence, by (V2′), $r(\tilde{x}, i, j, i, l) \geq 0$.

Upon considering all i, we have a finite number of values for l, namely the $l_{i,j,k}$; picking $l = \max\{l_{i,j,k}\}$ gives an l such that $D\tilde{f}(\tilde{x}) + lI : K \mapsto K, \forall\, \tilde{x} \in N$. □

To show that (V2′) $\not\Rightarrow$ (K1) in general, we present the following example.

Example 16: Consider the linear system $\dot{\tilde{x}} = \tilde{f}(\tilde{x}) = A\tilde{x}$, where

$$A = D\tilde{f} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We look at the right, circular cone K in $\mathbb{R}^3$ given by $z \geq \sqrt{x^2 + y^2}$. The inward normal at the point $(a, b, \sqrt{a^2 + b^2})$ on the boundary of the cone is $(-a, -b, \sqrt{a^2 + b^2})$; condition (V2′) is satisfied:

$$(D\tilde{f})\tilde{e} \cdot \tilde{N}_{\tilde{e}} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ \sqrt{a^2 + b^2} \end{pmatrix} \cdot \begin{pmatrix} -a \\ -b \\ \sqrt{a^2 + b^2} \end{pmatrix} = 0.$$

But, testing condition (K1) yields

$$(D\tilde{f}+lI)\begin{pmatrix} a \\ b \\ \sqrt{a^2+b^2} \end{pmatrix} = \begin{pmatrix} l & 1 & 0 \\ -1 & l & 0 \\ 0 & 0 & l \end{pmatrix}\begin{pmatrix} a \\ b \\ \sqrt{a^2+b^2} \end{pmatrix}$$
$$= \begin{pmatrix} b \\ -a \\ 0 \end{pmatrix} + l\begin{pmatrix} a \\ b \\ \sqrt{a^2+b^2} \end{pmatrix},$$

which will not be in K for any real l. To see this, realize that A projects vectors onto the $xy-$plane and rotates vectors by $\frac{\pi}{2}$ in the $xy-$plane; adding any amount of a boundary vector to its image under A cannot push the image back into K. Algebraically, we can check the resultant vector in the inequality that defines the cone:

$$z = l\sqrt{a^2+b^2} \not\geq (1+l^2)\sqrt{a^2+b^2} = \sqrt{x^2+y^2}.$$

In this case, (V2′) is satisfied, but (K1) is not. [37] considers this same example as *Beispiel* 5.

Finally, we observe that the condition (V2′) is in fact a necessary condition for order preserving flows with respect to a closed, convex cone.

Theorem 31 *If ϕ is monotone with respect to a closed, convex cone K then $\exists$ N compact such that, $\forall\ \tilde{x} \in N$, $(D\tilde{f}(\tilde{x}))\tilde{e}\cdot\tilde{N}_{\tilde{e}} \geq 0$, $\forall\ \tilde{e} \in relbdy(K)$, where $\tilde{e}\cdot\tilde{N}_{\tilde{e}} = 0$, $\tilde{N}_{\tilde{e}} \in K^*$.*

Proof: If ϕ is monotone with respect to a closed, convex cone K, then for $\tilde{e}$ a unit vector in K and $\epsilon > 0$

$$\phi_t(\tilde{x}_0+\epsilon\tilde{e}) \geq_K \phi_t(\tilde{x}_0) \Rightarrow \frac{\phi_t(\tilde{x}_0+\epsilon\tilde{e})-\phi_t(\tilde{x}_0)}{\epsilon}\cdot\tilde{N}_{\tilde{e}} \geq 0.$$

Letting $\epsilon \to 0$ gives

$$(D_{\tilde{x}_0}\phi_t)\tilde{e} \cdot \tilde{N}_{\tilde{e}} \geq 0.$$

Let $\psi(t) = (D_{\tilde{x}_0}\phi_t)\tilde{e} \cdot \tilde{N}_{\tilde{e}}$; then $\psi(t) \geq 0$ for $t \geq 0$ and $\psi(0) = 0$. We conclude that $\dot{\psi}(0) \geq 0$. Now

$$\dot{\psi}(t) = D\tilde{f}(\phi_t)D_{\tilde{x}_0}\phi_t\tilde{e} \cdot \tilde{N}_{\tilde{e}},$$

giving

$$\dot{\psi}(0) = D\tilde{f}(\tilde{x}_0)\tilde{e} \cdot \tilde{N}_{\tilde{e}},$$

and the desired result follows. □

From Theorem 30, we know that (V2′) and (K1) are equivalent for polyhedral cones, so (K1) is a necessary and sufficient condition for a monotone flow with respect to a polyhedral cone.

One might ask whether the condition for strong monotonicity, namely that there exists in addition a positive integer m such that

$$\left(D\tilde{f}(\tilde{x}) + lI\right)^m : K \setminus \{\tilde{0}\} \mapsto relint(K),\ \forall\, \tilde{x} \in N,$$

is also necessary for polyhedral cones or what the analogous condition in the form of (V2′) is. This final question remains unanswered. The following example shows that this condition for strong monotonicity is not necessary.

Example 17: Suppose that our system $\dot{\tilde{x}} = \tilde{f}(\tilde{x})$ is given by

$$\begin{aligned} \dot{x} &= y^3, \text{and} \\ \dot{y} &= x^3; \end{aligned}$$

then

$$D\tilde{f} = \begin{pmatrix} 0 & 3y^2 \\ 3x^2 & 0 \end{pmatrix}$$

satisfies $D\tilde{f} : \bar{\mathcal{O}} \mapsto \bar{\mathcal{O}}$. This gives a monotone flow with respect to the orthant. However, for any neighbourhood N containing the origin, we have for $\tilde{x} = \tilde{0}$,

$$\left(D\tilde{f}(\tilde{x}) + lI\right)^m : K \setminus \{\tilde{0}\} \not\mapsto relint(K),$$

so the condition for strong monotonicity fails. Note that solutions with initial value away from the origin are positive for all positive. Using Corollary 12 of Chapter 2, we see that the directed multigraph for this example consists of a single directed cycle with two positive edges. This system induces a strongly monotone flow even though our condition is not satisfied.

3.2.2 A Graph Theoretic Approach

Theorem 26 gives strong monotonicity for super- and sub-solutions $\tilde{w}(t)$ and $\tilde{v}(t)$ relative to a closed, convex cone K provided that, for any compact set N, $\exists\ l$ and m such that

$$D\tilde{f}(\tilde{x}) + lI \ : \ K \mapsto K,\ \forall\ \tilde{x} \in N, \text{ and} \tag{3.33}$$

$$\left(D\tilde{f}(\tilde{x}) + lI\right)^m \ : \ K \setminus \{\tilde{0}\} \mapsto relint(K),\ \forall\ \tilde{x} \in N. \tag{3.34}$$

As remarked upon earlier, (3.34) will only be able to hold for pointed cones. Based on the case of the orthant, we might expect that strong monotonicity has something to do with irreducibility of the matrix $D\tilde{f}(\tilde{x}) + lI$. Irreducibility with respect to a proper cone is discussed in [5]; to avoid confusion, the terminology K-irreducible is used. [5] presents the following equivalences for an $n \times n$ matrix A:

(i) $A : K \mapsto K$ is K-irreducible,

(ii) No eigenvectors of A are on the boundary of K,

(iii) Only the trivial faces of K are left invariant by A, and

(iv) $(A+I)^{n-1} : K \setminus \{\tilde{0}\} \mapsto relint(K)$,

where, in each case, K is a proper cone. The final item on this list is of immediate interest; it is closely related to (3.34). Rewrite (3.33) and (3.34) with l replaced by $l-1$ and let $A = D\tilde{f}(\tilde{x}) + (l-1)I$; then (3.34) implies (iv) above with $m = n-1$. So, Theorem 26 achieves strong monotonicity with respect to proper cones by demanding that $D\tilde{f}(\tilde{x}) + (l-1)I$ be K-irreducible. We could formulate our conditions as follows: for any compact set N, $\exists\ l$ and m such that

$$D\tilde{f}(\tilde{x}) + lI \ : \ K \mapsto K,\ \forall\ \tilde{x} \in N, \text{ and} \tag{3.35}$$

$$\left(D\tilde{f}(\tilde{x}) + (l+1)I\right)^m \ : \ K \setminus \{\tilde{0}\} \mapsto relint(K),\ \forall\ \tilde{x} \in N. \tag{3.36}$$

Of course, by picking l large enough, both (3.33) and (3.34) can be satisfied for the same l.

As in the case of the orthant, we can present the strict sign condition (3.34) graph theoretically. (3.34) requires that K be closed, convex, and pointed. In fact, the upcoming results will depend on whether K has the same number of generators as its dimension; when this is the case, we will say that K is *n-generated*. n-generated cones have a useful property: the non-negative span of any subset of generators yields a face of K.

We will use a graph based on the faces of the cone. For $\tilde{k} \in K$, define

$$F_{\tilde{k}} = \{\tilde{w} \in K : \exists\ \alpha > 0 \text{ such that } \alpha\tilde{w} \leq_K \tilde{k}\}.$$

$F_{\tilde{k}}$ is the smallest face containing $\tilde{k} \in K$. For a pointed, polyhedral cone satisfying (3.33) with generators labelled $\tilde{e}_i$, $i = \{1, \ldots, n_k\}$, we can construct a directed multigraph $G_{K,p}(\tilde{f}, N)$, where p is a positive integer, on the vertices $\{g_1, \ldots, g_{n_k}\}$

as follows. For each i, let $\tilde{k}_{p,i} = (D\tilde{f}(\tilde{x}) + (l+1)I)^p \tilde{e}_i$. Draw a directed edge from g_i to g_j, $i \neq j$, if $\tilde{e}_j \in F_{\tilde{k}_{p,i}}$, $\forall\, \tilde{x} \in N$. The following theorem gives a graph theoretic condition for K-irreducibility.

Theorem 32 *Suppose that $\tilde{f}(\tilde{x})$ is continuously differentiable in $\tilde{x}$ on compact subsets of $\mathbf{R}^n$ and that for any compact set, N, $\exists\ l = l(N)$ such that*

$$D\tilde{f}(\tilde{x}) + lI : K \mapsto K,\ \forall\, \tilde{x} \in N,$$

where K is a pointed, polyhedral cone; then $\forall$ compact N

$$D\tilde{f}(\tilde{x}) + lI \text{ is } K\text{-irreducible } \forall\, \tilde{x} \in N \Leftarrow G_{K,1}(\tilde{f}, N) \text{ is strongly connected.}$$

If, in addition, K is n-generated then

$$D\tilde{f}(\tilde{x}) + lI \text{ is } K\text{-irreducible } \forall\, \tilde{x} \in N \Rightarrow G_{K,1}(\tilde{f}, N) \text{ is strongly connected.}$$

Proof: ($\Leftarrow$) We prove the contrapositive. Suppose that $D\tilde{f}(\tilde{x}) + lI$ is K-reducible; then some nontrivial face F of K is left invariant by it. Let $\tilde{e}_i$, $i = \{1, \ldots, n_F\}$, be the generators of F, where $0 < n_F < n$. The strongly connected subgraph on the vertices $\{g_1, \ldots, g_{n_F}\}$ has no outward edges. $G_{K,1}(\tilde{f}, N)$ is not strongly connected.

($\Rightarrow$, K n-generated) We prove the contrapositive. Suppose $G_{K,1}(\tilde{f}, N)$ is not strongly connected; then there exists a strongly connected subgraph on the vertices $\{g_1, \ldots, g_{n_1}\}$, $1 \geq n_1 < n$, which is not strongly connected to the remainder of G. From the rules of construction and because K is n-generated, this means that $D\tilde{f}(\tilde{x}) + (l+1)I : F_{\tilde{k}_{1,i}} \mapsto F_{\tilde{k}_{1,i}}$, for each $i = \{1, \ldots, n_1\}$; thus, $D\tilde{f}(\tilde{x}) + (l+1)I$ and, hence, $D\tilde{f}(\tilde{x}) + lI$ leave at least one nontrivial face of K invariant and must be K-reducible. □

Remark: The graphs in the examples will not change based on N.

We provide the following example from [2] to show that the forward direction of the theorem does not hold when K is not n-generated.

Example 18: Let K be the proper cone in $\mathbb{R}^4$ with generators

$$\tilde{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tilde{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tilde{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \tilde{e}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \text{ and } \tilde{e}_5 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

Let $D\tilde{f}$ be the projection onto the $x_1x_2x_3$ subspace following by a linear mapping on the range of the projection given by $\tilde{e}_1 \mapsto \tilde{e}_1+\tilde{e}_2$, $\tilde{e}_2 \mapsto \tilde{e}_2+\tilde{e}_3$, and $\tilde{e}_3 \mapsto \tilde{e}_3+\tilde{e}_1$; then $D\tilde{f}K \subset sp^+\{\tilde{e}_1, \tilde{e}_2, \tilde{e}_3\}$. Better yet, $(D\tilde{f})^2\tilde{e}_i \in int(K)$, $\forall\, i$, so $D\tilde{f}$ is K-irreducible. However, $(D\tilde{f}+I)\tilde{e}_1 = 2\tilde{e}_1+\tilde{e}_2$, $(D\tilde{f}+I)\tilde{e}_2 = 2\tilde{e}_2+\tilde{e}_3$, and $(D\tilde{f}+I)\tilde{e}_3 = 2\tilde{e}_3+\tilde{e}_1$. In $G_{K,1}(\tilde{f}, N)$, there is no path from g_1 to g_4 say; the graph is not strongly connected.

Most systems will not be irreducible; it is desirable to have a result which gives us partial strong monotonicity with respect to convex cones. Lemma 34 leads to the primary result of this section. We will need the following lemma to establish Lemma 34.

Lemma 33 *Let $\tilde{e}_1, \cdots, \tilde{e}_n$ be the generators of a closed, convex, pointed cone K. Suppose that $\tilde{e}_1, \cdots, \tilde{e}_m$, $m \leq n$, generate a face of K. Then*

$$F_{\tilde{v}} = sp^+\{\tilde{e}_1, \cdots, \tilde{e}_m\} \iff \tilde{v} = \sum_{i=1}^{m} c_i\tilde{e}_i,\ c_i > 0.$$

Proof: ($\Rightarrow$) If $F_{\tilde{v}} = sp^+\{\tilde{e}_1, \cdots, \tilde{e}_m\}$, then the finite list of ways of expressing $\tilde{v}$ as non-negative combinations of the $\tilde{e}_i$ must include each $\tilde{e}_i$, $1 \leq i \leq m$, in some

combination. If this is the not the case, then $F_{\tilde{v}}$ is not the smallest face containing $\tilde{v}$. Adding all of the expressions for $\tilde{v}$ and dividing by the total number of such expressions gives $\tilde{v}$ as a positive combination of all of the $\tilde{e}_i$.

($\Leftarrow$) By definition, $\tilde{w}$ is in $F_{\tilde{v}}$ if there is a positive α such that $\tilde{v} - \alpha\tilde{w} \in K$. To see which generators of K are in $F_{\tilde{v}}$, notice that

$$\tilde{v} - \alpha\tilde{e}_j = \sum_{i=1}^{m} c_i\tilde{e}_i - \alpha\tilde{e}_j,\ c_i > 0.$$

It is clear that only $\tilde{e}_1, \cdots, \tilde{e}_m$ are in $F_{\tilde{v}}$. The result follows. □

Lemma 34 *For K a pointed, polyhedral cone, if the ordered vertex pair (g_i, g_j) is strongly connected in $G_{K,1}(\tilde{f}, N)$ then (g_i, g_j) is strongly connected in $G_{K,p}(\tilde{f}, N)$, $\forall\ p \geq 1$. If K is n-generated then the converse holds.*

Proof: The proof is by induction on p. The result is true for $p = 1$. Assume that the result holds for a particular $p > 1$. We consider the case $p + 1$. Since the result holds for p, we need only show that the ordered vertex pair (g_i, g_j) is strongly connected in $G_{K,p}(\tilde{f}, N)$ if and only if (g_i, g_j) is strongly connected in $G_{K,p+1}(\tilde{f}, N)$. Notice that

$$\tilde{k}_{p+1,i} = (D\tilde{f}(\tilde{x}) + (l+1)I)^{p+1}\tilde{e}_i = (D\tilde{f}(\tilde{x}) + (l+1)I)\tilde{k}_{p,i}.$$

Suppose the generators of $F_{\tilde{k}_{p,i}}$ are labelled $\{\tilde{e}_{r_1}, \ldots, \tilde{e}_{r_s}, \tilde{e}_i\}$; notice that $\tilde{e}_i$ must be one of them. This gives, using Lemma 33,

$$\text{generators of } F_{\tilde{k}_{p+1,i}} \supseteq \text{generators of } F_{\tilde{k}_{p,r_1}} \cup \text{generators of } F_{\tilde{k}_{p,r_2}} \cup \cdots \cup \text{generators of } F_{\tilde{k}_{p,r_s}} \cup \text{generators of } F_{\tilde{k}_{p,i}}; \quad (3.37)$$

furthermore, in the multigraph $G_{K,p}(\tilde{f}, N)$, there is a directed edge from g_i to each g_{r_s}.

($\Rightarrow$) Using (3.37), if $\tilde{e}_j \in F_{\tilde{k}_{p,i}}$, then $\tilde{e}_j \in F_{\tilde{k}_{p+1,i}}$. In fact, this means that all edges in $G_{K,p}(\tilde{f}, N)$ exist in $G_{K,p+1}(\tilde{f}, N)$ as well; if there is a directed path from g_i to g_j in $G_{K,p}(\tilde{f}, N)$, that same path exists in $G_{K,p+1}(\tilde{f}, N)$. The claim follows.

($\Leftarrow$, K n-generated) In this case, equality holds in (3.37). We use induction on the length of the shortest directed path from g_i to g_j in $G_{K,p+1}(\tilde{f}, N)$. Suppose the shortest directed path from g_i to g_j in $G_{K,p+1}(\tilde{f}, N)$ has length one, that there is a directed edge from g_i to g_j in $G_{K,p+1}(\tilde{f}, N)$. This means that $\tilde{e}_j \in F_{\tilde{k}_{p+1,i}}$. Using (3.37), either $\tilde{e}_j \in F_{\tilde{k}_{p,i}}$ or $\tilde{e}_j \in F_{\tilde{k}_{p,r_s}}$ for some s. In the first case, we are done. In the second case, there is a directed path of length two in $G_{K,p}(\tilde{f}, N)$ from g_i to g_j, passing through g_{r_s}. We assume the claim holds for a shortest directed path of length b. Consider a shortest directed path of length $b+1$ on the vertices $\tilde{g}^i, \tilde{g}^{q_1}, \ldots, \tilde{g}^{q_b}, \tilde{g}^j$, in this order. To avoid a simple contradiction this means that the shortest directed path from $\tilde{g}^{q_1}$ to $\tilde{g}^j$ in $G_{K,p+1}(\tilde{f}, N)$ has length b. The claim holds for this directed path by assumption: there is a directed path from g_{q_1} to g_j in $G_{K,p}(\tilde{f}, N)$. We need to show that there is a directed path from g_i to g_{q_1} in $G_{K,p}(\tilde{f}, N)$. Since $\tilde{e}_{q_1} \in F_{\tilde{k}_{p+1,i}}$, using (3.37), either $\tilde{e}_{q_1} \in F_{\tilde{k}_{p,i}}$ or $\tilde{e}_{q_1} \in F_{\tilde{k}_{p,r_s}}$ for some s. In the first case, the edge exists in $G_{K,p}(\tilde{f}, N)$ and we are done. In the second case, there is a directed path of length two in $G_{K,p}(\tilde{f}, N)$ from g_i to g_{q_1}, passing through g_{r_s}, and we are done. The claim follows. The lemma is proved. □

The following example shows that the converse of Lemma 34 does not hold when K is not n-generated.

Example 19: Let K be the proper cone in $\mathbf{R}^3$ with four generators, $\tilde{e}_1$,$\tilde{e}_2$,$\tilde{e}_3$, and $\tilde{e}_4$, where any three of the four generators are linearly independent and $\tilde{e}_1+\tilde{e}_3 = \tilde{e}_2+\tilde{e}_4$. Suppose that $D\tilde{f} : \tilde{e}_1 \mapsto \tilde{e}_1 + \tilde{e}_2$, $D\tilde{f} : \tilde{e}_2 \mapsto 3\tilde{e}_2 + \tilde{e}_3$, and $D\tilde{f} : \tilde{e}_3 \mapsto \tilde{e}_2 + 2\tilde{e}_3$. It is then easy to check that $D\tilde{f} : \tilde{e}_4 \mapsto \tilde{e}_4$. $G_{K,1}(\tilde{f}, N)$ and $G_{K,2}(\tilde{f}, N)$ are presented in Figure 22; $G_{K,2}(\tilde{f}, N)$ contains a connection which $G_{K,1}(\tilde{f}, N)$ does not contain.

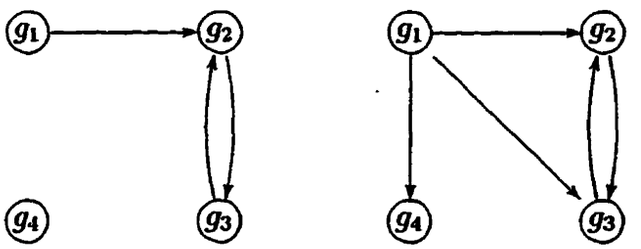


Figure 22: $G_{K,1}(\tilde{f}, N)$ and $G_{K,2}(\tilde{f}, N)$ for Example 19.

The graph theoretic approach offers an advantage over the linear algebra approach insofar as one need not calculate powers of the matrix $D\tilde{f}(\tilde{x}) + (l+1)I$. Consider the following proposition:

Proposition 35 *$G_{K,p}(\tilde{f}, N)$, $p > 1$, can be constructed by using $G_{K,1}(\tilde{f}, N)$ and the face structure of K.*

Remark: We offer no careful proof of Proposition 35. Intuitively, however, knowing the face structure of K and knowing in which smallest face of K the image of each generator of K under the matrix $D\tilde{f}(\tilde{x}) + (l+1)I$ lies allows us to determine where each generator is mapped by higher powers of the matrix. The graphs can be generated inductively from $G_{K,1}(\tilde{f}, N)$.

We arrive at the following result, which will lead us to the result we will use in the examples.

Theorem 36 *Suppose that $\tilde{f}(\tilde{x})$ is continuously differentiable in $\tilde{x}$ on compact subsets of $\mathbf{R}^n$ and that for any compact set, N, $\exists\ l = l(N)$ such that*

$$D\tilde{f}(\tilde{x}) + lI : K \mapsto K,\ \forall\ \tilde{x} \in N,$$

where K is a pointed, polyhedral cone in $\mathbf{R}^n$. Further assume that for N and l chosen as above,

$\forall\ \tilde{e}_i \in F_{\tilde{k}}, g_i$ is strongly connected to all other vertices in $G_{K,n-1}(\tilde{f}, N), \forall\ \tilde{x} \in N$,

for some vector $\tilde{k} \in K \setminus \{\tilde{0}\}$*; then*

$$\frac{\partial \tilde{x}}{\partial \tilde{k}} >_K \tilde{0},$$

for $t > 0$ ($t \geq 0$ *if* $\tilde{k} >_K \tilde{0}$).

Corollary 37 will prove particularly useful in applications; we suppose that the key hypotheses apply to the orthant (where solutions for chemical kinetics or epidemiological problems live), eliminating the need to check all compact sets. In this case, the multigraph $G_{K,1}(\tilde{f}, N)$ is replaced by $G_{K,1}(\tilde{f}, \mathcal{O})$, where we are assuming that the graph has the same structure at all points of the positive orthant.

Corollary 37 *Assume that* $\bar{\mathcal{O}}$ *is positively invariant and that the positivity assumption holds* ($\tilde{x}(t) > \tilde{0}$ *for* $t > 0$). *Suppose that* $\tilde{f}(\tilde{x})$ *is continuously differentiable on* $\bar{\mathcal{O}}$ *and that* $\exists\ l$ *such that*

$$D\tilde{f}(\tilde{x}) + lI : K \mapsto K,\ \forall\ \tilde{x} \in \bar{\mathcal{O}},$$

where K *is a pointed, polyhedral cone in* $\mathbf{R}^n$. *Further assume that for* l *chosen as above,*

$$\forall\ \tilde{e}_i \in F_{\tilde{k}}, g_i \text{ is strongly connected to all other vertices in } G_{K,1}(\tilde{f}, \mathcal{O}),$$

for some vector $\tilde{k} \in K \setminus \{\tilde{0}\}$*; then*

$$\frac{\partial \tilde{x}}{\partial \tilde{k}} >_K \tilde{0},$$

for $t > 0$ ($t \geq 0$ *if* $\tilde{k} >_K \tilde{0}$).

At first glance, it may seem that Theorem 36 demands more than its analog in Chapter 2, since requiring that $D\tilde{f}$ preserves K demands that the system induces

an order preserving flow with respect to K. We had results in Chapter 2 guaranteeing monotonicity in some components when the system does not induce an order preserving flow with respect to an orthant. For the partial monotonicity results of Chapter 2, it is the case that there is a simple polyhedral cone with respect to which the flow is order preserving. Using the hypotheses of Corollary 12, we state the following result.

Theorem 38 *If (v_i, v_j) is positively (negatively) consistently strongly connected in $G(\tilde{f}, \Omega)$, then the system is order preserving with respect to a polyhedral cone with generators that are standard basis vectors or the negative of standard basis vectors.*

Proof: Without loss of generality, use $i = 1$ and, as in the proof of Theorem 9, partition the system by defining the disjoint sets

$$\begin{aligned}
\mathcal{Q}_1 &= \{v_k : (v_1, v_k) \text{ is positively consistently strongly connected in } G(\tilde{f}, \Omega)\},\\
\mathcal{Q}_2 &= \{v_k : (v_1, v_k) \text{ is negatively consistently strongly connected in } G(\tilde{f}, \Omega)\},\\
\mathcal{R} &= \{v_k : (v_1, v_k) \text{ is not strongly connected in } G(\tilde{f}, \Omega)\}, \text{ and}\\
\mathcal{S} &= \{v_k : (v_1, v_k) \text{ is inconsistently strongly connected in } G(\tilde{f}, \Omega)\}.
\end{aligned}$$

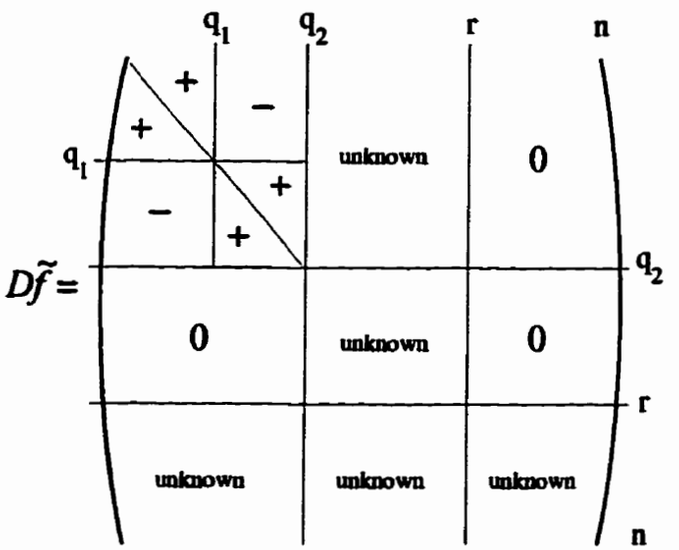
Relabel the vertices so that

$$\begin{aligned}
v_1, \ldots, v_{q_1} &\in \mathcal{Q}_1,\\
v_{q_1+1}, \ldots, v_{q_2} &\in \mathcal{Q}_2,\\
v_{q_2+1}, \ldots, v_r &\in \mathcal{R}, \text{ and}\\
v_{r+1}, \ldots, v_n &\in \mathcal{S}.
\end{aligned}$$

With the corresponding relabelling of x_i, $1 \leq i \leq n$, the system (2.7) takes on the form

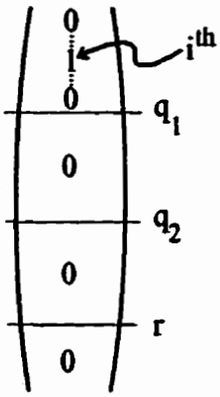
$$\begin{aligned} \dot{x}_i &= f_i(x_1, \ldots, x_r), & 1 \leq i \leq q_2, \\ \dot{x}_i &= f_i(x_{q_2+1}, \ldots, x_r), & q_2 + 1 \leq i \leq r, \\ \dot{x}_i &= f_i(x_1, \ldots, x_n), & r+1 \leq i \leq n. \end{aligned}$$

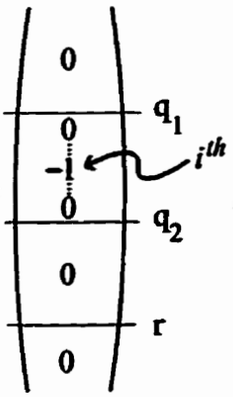
In this setting the Jacobian matrix has the form



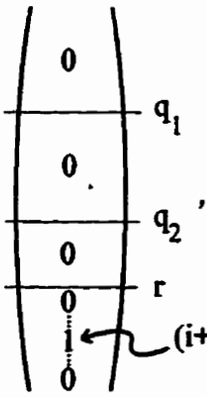
where the '+' signs represent non-negative entries, the '−' signs represent non-positive entries, the zeroes represent entries which are identically zero, and 'unknown' is written in blocks for which we do not have sign information. The polyhedral cone K preserved by this $D\tilde{f}$ has extreme rays $\tilde{e}_i$, $i = 1, \ldots, q_2 + 2(n - r + 1)$,

each +1 or −1 times a standard basis vector, where

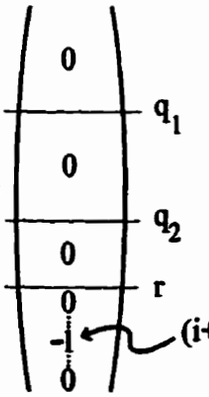
$\tilde{e}_i =$ 

$, 1 \leq i \leq q_1,$ $\tilde{e}_i =$ 

$, q_1 < i \leq q_2,$

$\tilde{e}_i =$ 

$, q_2 < i \leq q_2+n-r+1,$

$\tilde{e}_i =$ 

$, q_2+n-r+1 < i \leq q_2+2(n-r+1).$

It is easy to check that this cone is preserved by $D\tilde{f}$:

$$(D\tilde{f})\tilde{e}_i = \begin{cases} \sum_{i=1}^{q_2+2(n-r+1)} c_i\tilde{e}_i, & c_i \geq 0, \quad 1 \leq i \leq q_2 \\ \sum_{i=q_2+1}^{q_2+2(n-r+1)} c_i\tilde{e}_i, & c_i \geq 0, \quad q_2+1 \leq i \leq q_2+2(n-r+1) \end{cases};$$

all images of extreme vectors are in K. □

By Theorem 38, supposing that our system does induce some monotonicity, assuming that there exists a convex cone K satisfying (3.33) places no additional restrictions on our system.

Remark: In Corollary 12, we concluded that the partial derivative of x_j with respect to x_i was of strict sign. This follows in Theorem 38 as well. With the vertices relabelled as in the proof of Theorem 38 and K so defined, $(D\tilde{f}(\tilde{x})+(l+1)I)^{q_2-1}\tilde{e}_1 \in$ *relint*(K). As in the proof of Corollary 29 with $\tilde{k} = \tilde{e}_1$, it follows that the partial derivative of x_j, $1 \leq j \leq q_2$, with respect to x_1 (the relabelled x_i) is of strict sign.

Example 20: We reconsider the problem of Example 14, the chemical reaction mechanism

$$\alpha_1 A_1 + \alpha_2 A_2 \xrightarrow{k} \bar{\alpha}_1 A_1 + \bar{\alpha}_2 A_2, \tag{3.38}$$

with $\bar{\alpha}_1 > \alpha_1$ and $\alpha_2 > \bar{\alpha}_2$. The Jacobian matrix for this problem was

$$D\tilde{f} = k(A_1(t))^{\alpha_1-1}(A_2(t))^{\alpha_2-1}\begin{bmatrix} \alpha_1(\bar{\alpha}_1-\alpha_1)A_2(t) & \alpha_2(\bar{\alpha}_1-\alpha_1)A_1(t) \\ \alpha_1(\bar{\alpha}_2-\alpha_2)A_2(t) & \alpha_2(\bar{\alpha}_2-\alpha_2)A_1(t) \end{bmatrix}. \tag{3.39}$$

The proper cone used in that example had extreme rays

$$\tilde{a} = \begin{bmatrix} 1 \\ \mu_2 \end{bmatrix} \text{ and } \tilde{b} = \tilde{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where

$$\mu_2 = \frac{\bar{\alpha}_2 - \alpha_2}{\bar{\alpha}_1 - \alpha_1}.$$

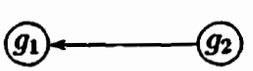
From the work of Example 12, we already know that

$$\frac{\partial A_1}{\partial A_{10}}(t) \geq 0 \text{ and } \frac{\partial A_2}{\partial A_{10}}(t) \leq 0, \ \forall\, t \geq 0; \tag{3.40}$$

Let us see if the inequalities are strict. We will construct a graph on two vertices, g_1 corresponding to $\tilde{a}$ and g_2 corresponding to $\tilde{b}$. Now,

$$\begin{aligned}
(D\tilde{f})\tilde{a} &= k(A_1(t))^{\alpha_1-1}(A_2(t))^{\alpha_2-1}\begin{bmatrix} \alpha_1(\bar{\alpha}_1-\alpha_1)A_2\left[1+\frac{\alpha_2 A_1}{\alpha_1 A_2}\mu_2\right] \\ \alpha_1(\bar{\alpha}_2-\alpha_2)A_2\left[1+\frac{\alpha_2 A_1}{\alpha_1 A_2}\mu_2\right] \end{bmatrix} \\
&= \alpha_1(\bar{\alpha}_1-\alpha_1)k(A_1(t))^{\alpha_1-1}(A_2(t))^{\alpha_2}\left[1+\frac{\alpha_2 A_1}{\alpha_1 A_2}\mu_2\right]\begin{bmatrix} 1 \\ \mu_2 \end{bmatrix} \\
&= \text{a multiple of } \tilde{a}, \\
(D\tilde{f})\tilde{b} &= k(A_1(t))^{\alpha_1-1}(A_2(t))^{\alpha_2-1}\begin{bmatrix} \alpha_1(\bar{\alpha}_1-\alpha_1)A_2 \\ \alpha_1(\bar{\alpha}_2-\alpha_2)A_2 \end{bmatrix} \\
&= \alpha_1(\bar{\alpha}_1-\alpha_1)k(A_1(t))^{\alpha_1-1}(A_2(t))^{\alpha_2}\begin{bmatrix} 1 \\ \mu_2 \end{bmatrix} \\
&= \text{a positive multiple of } \tilde{a};
\end{aligned}$$

hence, $D\tilde{f}+lI : \tilde{a} \mapsto l_1\tilde{a}$, $l_1 > 0$, and $D\tilde{f}+lI : \tilde{b} \mapsto relint(K)$ for $l > 0$. Our graph has but a single edge:



Since g_2 corresponds to extreme ray $\tilde{e}_1$, we can conclude by Theorem 36 that the partial derivatives in (3.40) have strict sign for $t > 0$.

3.2.3 Finding Cones

Given a particular system (2.7), how does one find a cone K which satisfies the conditions of the theorems in this chapter? This is a very difficult question. The

majority of the literature in this area is abstract with no practical focus. An extremely useful contribution was made by J. Vandergraft in [36]. The backbone of this work is Perron-Frobenius theory. We present some results from this paper.

Theorem 39 *If K is a solid cone and $A\tilde{x} \in K$, $\forall\ \tilde{x} \in K$, then*

(i) $\rho(A)$, the spectral radius of A, is an eigenvalue;

(ii) The degree of $\rho(A)$ is no smaller than the degree of any other eigenvalue having the same modulus; and

(iii) K contains an eigenvector corresponding to $\rho(A)$.

Furthermore, conditions (i) and (ii) are sufficient to insure that A leaves invariant a solid cone.

Theorem 40 *A satisfying $A\tilde{x} \in K$, $\forall\ \tilde{x} \in K$, is K-irreducible for some solid cone K*

(i) if and only if no eigenvector of A lies on the boundary of K;

(ii) if and only if one eigenvector lies in the interior of K;

(iii) implies $\rho(A)$ is a simple eigenvalue, any other eigenvalue with the same modulus is also simple, there is an eigenvector corresponding to $\rho(A)$ in the interior of K, and no other eigenvector lies in K.

Both of the above theorems require that the cone K be solid. This means that the cone must be of the same dimension as the space in which it resides. As it turns out, we will often be able to obtain results from cones of lower dimension than their

space (for example, two-dimensional cones in $\mathbb{R}^3$). To find non-solid cones, first find invariant subspaces of A and then restrict to these subspaces. Of course, having found such a cone, one can rewrite the matrix A restricted to the subspace spanned by the cone and verify that the conditions of the above theorems are satisfied. We will see this in the examples.

Work on convex polytopes is applicable to the problem of finding the facial structure of polyhedral cones, which is required for the strong monotonicity results. There is work on algorithms for finding the convex hull of a set of points or the representation of convex polyhedra in terms of faces in [25] and [6]. [45] offers an advanced discussion of the theory of convex polytopes. The discussion of face structure includes program code which produces a minimal system of facet-defining inequalities from a set of vertices. Facets are the faces of a polytope of one lower dimension than the polytope.

The following result will prove useful in the examples.

Proposition 41 *Let $\tilde{v}_1$, $\tilde{v}_2$, $\tilde{v}_3$, and $\tilde{v}_4 \in \mathbb{R}^3$, with any three of the four vectors being linearly independent. Suppose that*

$$a_1\tilde{v}_1 + a_2\tilde{v}_2 + a_3\tilde{v}_3 + a_4\tilde{v}_4 = \tilde{0}, \tag{3.41}$$

where $a_1a_2a_3a_4 \neq 0$; then $\tilde{v}_1$, $\tilde{v}_2$, $\tilde{v}_3$, and $\tilde{v}_4$ generate a polyhedral cone if and only if two of a_1, a_2, a_3, and a_4 are positive and two are negative.

Proof: ($\Leftarrow$) The vectors v_i can be scaled and relabelled so that (3.41) takes the form

$$\tilde{w}_1 + \tilde{w}_2 = \tilde{w}_3 + \tilde{w}_4. \tag{3.42}$$

With this labelling, when the vectors generate a polyhedral cone, the two dimensional faces of the cone are $F_{1,3} = sp^+\{\tilde{w}_1, \tilde{w}_3\}$, $F_{1,4} = sp^+\{\tilde{w}_1, \tilde{w}_4\}$, $F_{2,3} =$

$sp^+\{\tilde{w}_2, \tilde{w}_3\}$, and $F_{2,4} = sp^+\{\tilde{w}_2, \tilde{w}_4\}$. The vectors generate a polyhedral cone when for each face the two vectors $\tilde{w}_i$ not in the face both lie on the same side of the plane containing the face. Suppose this is not the case: assume that $\tilde{w}_2$ and $\tilde{w}_4$ lie on opposite sides of $F_{1,3}$. The line segment connecting $\tilde{w}_2$ and $\tilde{w}_4$ must intersect the plane spanned by $\tilde{w}_1$ and $\tilde{w}_3$. Mathematically, for some r, s, and t, $0 < t < 1$,

$$t\tilde{w}_2 + (1-t)\tilde{w}_4 = r\tilde{w}_1 + s\tilde{w}_3;$$

using (3.42) to eliminate $\tilde{w}_4$ gives

$$\tilde{w}_2 + (1-t)(\tilde{w}_1 - \tilde{w}_3) = r\tilde{w}_1 + s\tilde{w}_3.$$

Since any three of the four vectors are assumed to be linearly independent, this gives a contradiction.

($\Rightarrow$) Suppose that one a_i is negative and the others are positive; then, after scaling and relabelling, (3.41) gives

$$\tilde{w}_1 = \tilde{w}_2 + \tilde{w}_3 + \tilde{w}_4.$$

This means that $\tilde{w}_1$ is an interior vector of the polyhedral cone generated by $\tilde{w}_2$, $\tilde{w}_3$, and $\tilde{w}_4$, contradicting the assumption that all four vectors are extreme rays. The argument is similar if one a_i is positive and the others are negative.

Suppose that all four a_i are positive; then, after scaling and relabelling, (3.41) gives

$$\tilde{w}_1 + \tilde{w}_2 + \tilde{w}_3 + \tilde{w}_4 = \tilde{0}. \tag{3.43}$$

This suggests that $\tilde{0}$ is in the interior of the cone generated by the $\tilde{w}_i$, an unnerving implication. Notice that any vector can be expressed as a linear combination of three of the (linearly independent) $\tilde{w}_i$:

$$\tilde{x} = c_1\tilde{w}_1 + c_2\tilde{w}_2 + c_3\tilde{w}_3.$$

Using (3.43), we have for $\gamma > 0$

$$\begin{aligned}\tilde{x} &= c_1\tilde{w}_1 + c_2\tilde{w}_2 + c_3\tilde{w}_3 + \gamma(\tilde{w}_1 + \tilde{w}_2 + \tilde{w}_3 + \tilde{w}_4) \\ &= (c_1 + \gamma)\tilde{w}_1 + (c_2 + \gamma)\tilde{w}_2 + (c_3 + \gamma)\tilde{w}_3 + \gamma\tilde{w}_4;\end{aligned}$$

sure enough, any vector can be expressed as a positive combination of the $\tilde{w}_i$ for γ large enough. This means that any vector is in the interior of the cone generated by the $\tilde{w}_i$, a contradiction since the cone was assumed to be convex. The argument is similar if all four a_i are negative.

The only remaining possibility is that two a_i are negative and two are positive.

□

Notice that Proposition 41 also applies if we have four vectors in $\mathbb{R}^n$ satisfying the key hypotheses since we could simply work in the three-dimensional subspace spanned by the vectors.

3.3 Examples

When applying the graph theory of this chapter in the following examples, we will never use $G_{K,p}(\tilde{f}, \mathcal{O})$, $p > 1$. To simplify notation, we will use the label $G_K(\tilde{f}, \mathcal{O})$ for the case $p = 1$.

Example 21 (Chemical Kinetics): We return to Example 7, considering the chemical reaction

$$A_1 + A_2 \xrightarrow{k} A_3,$$

which leads to the system of differential equations, $\dot{\tilde{x}} = \tilde{f}(\tilde{x})$,

$$\dot{x}_1 = -kx_1x_2,$$

$$\dot{x}_2 = -kx_1x_2, \text{ and}$$
$$\dot{x}_3 = +kx_1x_2.$$

Using the machinery of the previous chapter, we could not determine the signs of the partial derivatives of $x_3(t)$ with respect to either $x_1(0)$ or $x_2(0)$. Let us try to apply the work of this chapter to this problem. We have

$$D\tilde{f} = \begin{pmatrix} -kx_2 & -kx_1 & 0 \\ -kx_2 & -kx_1 & 0 \\ kx_2 & kx_1 & 0 \end{pmatrix}.$$

This matrix $D\tilde{f} + lI$ has

eigenvalue l with corresponding eigenvectors $\begin{pmatrix} x_1 \\ -x_2 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, and

eigenvalue $l - k(x_1 + x_2)$ with corresponding eigenvector $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

For l sufficiently large, the maximal eigenvalue of $D\tilde{f}+lI$ is l; using Theorem 39, we know that any solid cone that is preserved by $D\tilde{f}$ must contain exactly one of the eigenvectors corresponding to eigenvalue l. Since we would like to draw conclusions on the signs of the partial derivatives of $x_3(t)$ with respect to either $x_1(0)$ or $x_2(0)$, our cone must include the x and y axes. After some thought (notice that the x and y axes are mapped by $D\tilde{f}$ to positive multiples of the vector $(-1,-1,1)^T$), we consider the proper (polyhedral) cone K with extreme rays

$$\tilde{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tilde{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \text{and} \tilde{e}_3 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix},$$

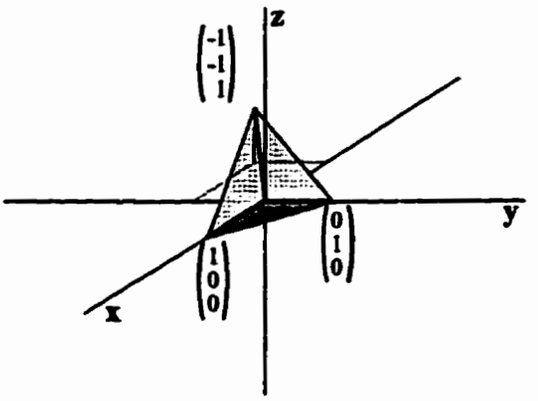


Figure 23: The proper (polyhedral) cone K for Example 21.

drawn in Figure 23. z is positive in K, which contains the x and y axes; hence, if K satisfies the essential hypothesis of Corollary 24 (if $D\tilde{f}$ preserves K), we will be able to conclude that the two derivatives of interest are both non-negative. We examine the images of the extreme rays of K under $D\tilde{f}$.

$$D\tilde{f}\begin{pmatrix}1\\0\\0\end{pmatrix} = kx_2\begin{pmatrix}-1\\-1\\1\end{pmatrix} \in K,$$

$$D\tilde{f}\begin{pmatrix}0\\1\\0\end{pmatrix} = kx_1\begin{pmatrix}-1\\-1\\1\end{pmatrix} \in K, \text{and}$$

$$D\tilde{f}\begin{pmatrix}-1\\-1\\1\end{pmatrix} = k(x_1+x_2)\begin{pmatrix}1\\1\\-1\end{pmatrix}$$

$$\Rightarrow \text{ for } \quad l \quad \text{sufficiently large }, (D\tilde{f}+lI)\begin{pmatrix}-1\\-1\\1\end{pmatrix} \in K.$$

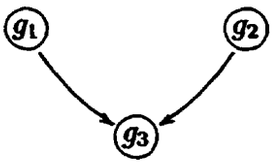


Figure 24: $G_K(\tilde{f}, \mathcal{O})$ for Example 21.

$D\tilde{f}$ preserves the cone K! By Corollary 24, we can conclude that for $t \geq 0$

$$\frac{\partial x_3}{\partial x_{1,0}} \geq 0 \text{ and } \frac{\partial x_3}{\partial x_{2,0}} \geq 0.$$

Now, if $D\tilde{f} + lI$ is K-irreducible, then the inequalities above are strict. We can use the graph theoretic approach of Corollary 37; we draw a graph on three vertices, with vertex g_1 (g_2,g_3) representing extreme ray $\tilde{e}_1$ ($\tilde{e}_2$,$\tilde{e}_3$). Figure 24 presents the graph $G_K(\tilde{f}, \mathcal{O})$ for this example. In this case, we can not apply Corollary 37, but all hope is not lost.

Notice that the image of all of the extreme rays under $D\tilde{f}$ lies on the line containing $\tilde{e}_3$. This means that the two dimensional cone K_1 (K_2) with extreme rays $\tilde{e}_1$ and $\tilde{e}_3$ ($\tilde{e}_2$ and $\tilde{e}_3$) is preserved by $D\tilde{f}$. The graph $G_{K_1}(\tilde{f}, \mathcal{O})$ ($G_{K_2}(\tilde{f}, \mathcal{O})$) consists of the subgraph of $G_K(\tilde{f}, \mathcal{O})$ on the vertices g_1 and g_3 (g_2 and g_3). Corollary 37 applies in each case, letting us conclude that earlier partial derivatives inequalities are strict for $t > 0$.

The cones K_1 and K_2 are not solid; hence, Theorem 39 and Theorem 40 do not apply. We can however construct a matrix representing the transformation $D\tilde{f}$ restricted to the two dimensional subspace corresponding to each cone; then the theorems will apply. We do this for K_1. The extreme rays are $\tilde{e}_1$ and $\tilde{e}_3$. The

images under $D\tilde{f}$ are

$$\begin{aligned} D\tilde{f}\begin{pmatrix}1\\0\\0\end{pmatrix} &= kx_1\begin{pmatrix}-1\\-1\\1\end{pmatrix} \\ &= kx_1\tilde{e}_3, \\ D\tilde{f}\begin{pmatrix}-1\\-1\\1\end{pmatrix} &= k(x_1+x_2)\begin{pmatrix}1\\1\\-1\end{pmatrix} \\ &= -k(x_1+x_2)\tilde{e}_3; \end{aligned}$$

hence, the transformation restricted to the two dimensional subspace spanned by $\tilde{e}_1$ and $\tilde{e}_3$ has the form

$$A = \begin{pmatrix} 0 & 0 \\ kx_1 & -k(x_1+x_2) \end{pmatrix}.$$

A has maximal eigenvalue 0 with corresponding eigenvector $[1,0]^T = \tilde{e}_1 \in K_1$. This agrees with the claims of Theorem 39; by Theorem 40, we conclude that the matrix is not K_1-irreducible, agreeing with the graph $G_{K_1}(\tilde{f},\mathcal{O})$, which is not strongly connected. A similar check of K_2 can be done.

Example 22 (Chemical Kinetics): Consider again the mechanism of Example 8, namely

$$A_1 + A_2 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} A_3.$$

This type of chain reaction was examined in [18]; elaborate arguments were required to established the monotonicity results presented in Table 8, where '++' means positive for $t \geq 0$, '+' means positive for $t > 0$, '0' means the derivative is zero for all time, and '*/+' means the derivative is positive for $x_{1,0} \leq x_{2,0}$, $\dot{x}_1 < 0$ and

	$x_1(t)$	$x_2(t)$	$x_3(t)$
$x_{1,0}$	++	−	*/+
$x_{2,0}$	−	++	*/+
$x_{3,0}$	+	+	++

Table 8: Signs of concentrations with respect to changes in initial concentrations for Example 22.

of both signs if $x_{1,0} > x_{2,0}$. This mechanism leads to the system of differential equations, $\dot{\tilde{x}} = \tilde{f}(\tilde{x})$,

$$\begin{aligned} \dot{x}_1 &= -k_1 x_1 x_2 + k_{-1} x_3, \\ \dot{x}_2 &= -k_1 x_1 x_2 + k_{-1} x_3, \text{ and} \\ \dot{x}_3 &= +k_1 x_1 x_2 - k_{-1} x_3, \end{aligned}$$

with Jacobian matrix

$$D\tilde{f} = \begin{pmatrix} -k_1 x_2 & -k_1 x_1 & k_{-1} \\ -k_1 x_2 & -k_1 x_1 & k_{-1} \\ k_1 x_2 & k_1 x_1 & -k_{-1} \end{pmatrix}.$$

The matrix $D\tilde{f} + lI$ has

eigenvalue l with corresponding eigenvectors $\begin{pmatrix} k_{-1} \\ 0 \\ k_1 x_2 \end{pmatrix}$ and $\begin{pmatrix} -k_1 x_1 \\ k_1 x_2 \\ 0 \end{pmatrix}$, and

eigenvalue $l - k_1 x_2 - k_1 x_1 - k_{-1}$ with corresponding eigenvector $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

Again, for l sufficiently large, l will be the maximal eigenvalue; hence, any cone that

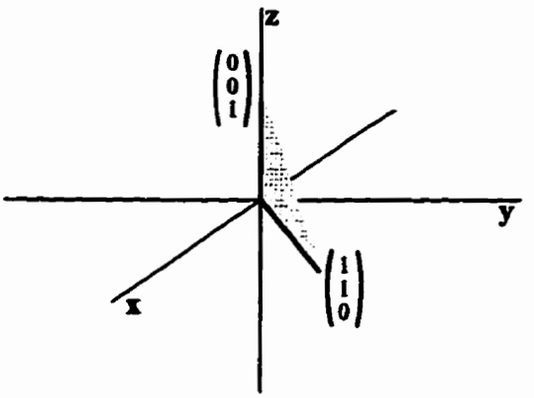


Figure 25: A proper (polyhedral) cone K_1 for Example 22.

we wish to use to establish some monotonicity results will have to contain exactly one of the corresponding eigenvectors in its interior.

The first cone we examine is the two dimensional proper (polyhedral) cone K_1 in $\mathbf{R}^3$ with extreme rays

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

pictured in Figure 25. The images of the extreme rays under $D\tilde{f}$ are easily calculated:

$$D\tilde{f}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = k_{-1}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \Rightarrow (D\tilde{f}+lI)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in K,\ l \geq k_1$$

$$D\tilde{f}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = k_1(x_1+x_2)\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \Rightarrow (D\tilde{f}+lI)\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \in K,\ l \text{ large enough.}$$

This cone is preserved by $D\tilde{f}$; we can conclude that, for $t \geq 0$,

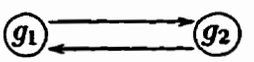
$$\frac{\partial x_1}{\partial x_{3,0}} \geq 0, \frac{\partial x_2}{\partial x_{3,0}} \geq 0, \text{ and } \frac{\partial x_3}{\partial x_{3,0}} \geq 0.$$

Using Corollary 37, we draw a graph on two vertices, with

$$\text{vertex } g_1 \text{ representing extreme ray } \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ and}$$

$$\text{vertex } g_2 \text{ representing extreme ray } \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix};$$

namely



Since the graph is strongly connected, we can conclude that the above partial derivatives are of strict sign for $t > 0$, giving us the bottom row of Table 8.

The second cone we examine is the three dimensional proper (polyhedral) cone K_2 in $\mathbf{R}^3$ with the four extreme rays

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix},$$

drawn in Figure 26. We calculate the images under $D\tilde{f}$ of the extreme rays.

$$D\tilde{f}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = k_1 x_2 \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix},$$

$$D\tilde{f}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = (k_1 x_2 + k_{-1}) \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix},$$

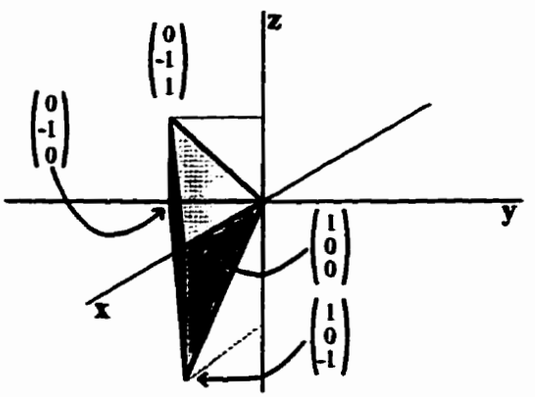


Figure 26: A proper (polyhedral) cone K_2 for Example 22.

$$D\tilde{f}\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = k_1 x_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \text{ and}$$

$$D\tilde{f}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = (k_1 x_1 + k_{-1}) \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

In each case, $D\tilde{f} + lI$ will map the extreme ray into the cone for l sufficiently large; the cone is preserved by $D\tilde{f}$. Considering the $x-$axis and negative $y-$axis, we can conclude that for $t \geq 0$

$$\frac{\partial x_1}{\partial x_{1,0}} \geq 0, \frac{\partial x_1}{\partial x_{2,0}} \leq 0, \frac{\partial x_2}{\partial x_{1,0}} \leq 0, \text{ and } \frac{\partial x_2}{\partial x_{2,0}} \geq 0.$$

Using Corollary 37 and the above calculations, assuming l is chosen sufficiently large, we would draw a graph on four vertices with each vertex connected to the other three. This strongly connected graph would tell us that the above partial derivatives are all of strict sign for $t > 0$. This gives the results in the upper left two-by-two block of Table 8.

Finally, we consider the three dimensional proper (polyhedral) cone K_3 in $\mathbf{R}^3$

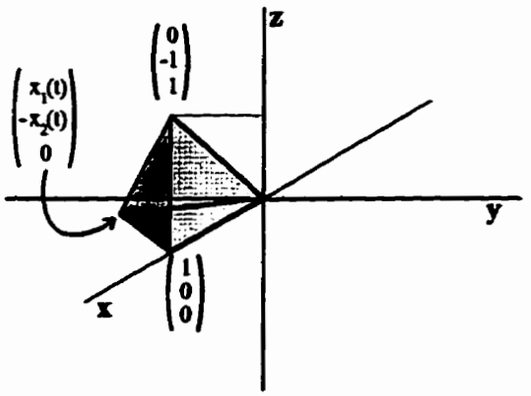


Figure 27: An expanding proper (polyhedral) cone K_3 for Example 22.

with the three extreme rays

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \text{ and } \begin{pmatrix} x_1(t) \\ -x_2(t) \\ 0 \end{pmatrix},$$

pictured in Figure 27. Notice that since $\dot{x}_1 = \dot{x}_2$, this cone is expanding if $\dot{x}_1 < 0$. Once again, we calculate the images under $D\tilde{f}$ of the extreme rays.

$$\begin{aligned} D\tilde{f}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} &= k_1 x_2 \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \\ D\tilde{f}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} &= (k_1 x_1 + k_{-1}) \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \text{ and} \\ D\tilde{f}\begin{pmatrix} x_1(t) \\ -x_2(t) \\ 0 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \end{aligned}$$

for l sufficiently large, $D\tilde{f} + lI$ maps each extreme ray into K_3. Since the cone is

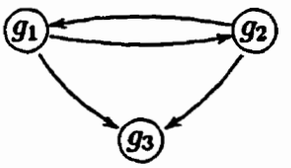


Figure 28: $G_{K_3}(\tilde{f}, \mathcal{O})$ for Example 22.

preserved by $D\tilde{f}$ we can conclude that

$$\frac{\partial x_3}{\partial x_{1,0}} \geq 0, t \geq 0,$$

and, by symmetry in x_1 and x_2,

$$\frac{\partial x_3}{\partial x_{2,0}} \geq 0, t \geq 0.$$

Using Corollary 37, we draw a graph on three vertices, with

$$\text{vertex } g_1 \text{ representing extreme ray } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

$$\text{vertex } g_2 \text{ representing extreme ray } \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}, \text{ and}$$

$$\text{vertex } g_3 \text{ representing extreme ray } \begin{pmatrix} x_1 \\ -x_2 \\ 0 \end{pmatrix};$$

the graph is presented in Figure 28. The graph is not strongly connected, so $D\tilde{f}+lI$ is not K-irreducible. But vertex g_1, corresponding to extreme ray $[1,0,0]^T$, is strongly connected to all other vertices; hence, by Corollary 37, assuming $\dot{x}_1 < 0$,

$$\frac{\partial x_3}{\partial x_{1,0}} > 0, t > 0,$$

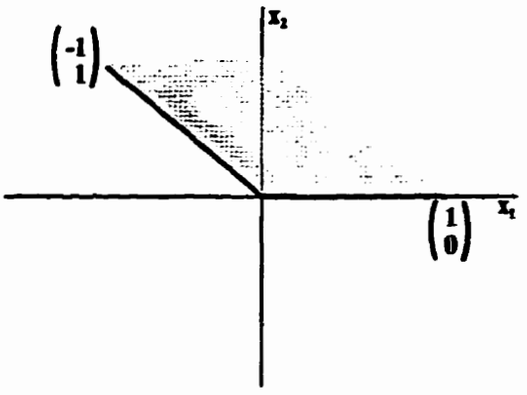


Figure 29: The proper (polyhedral) cone K_1 for Example 23.

and, by symmetry in x_1 and x_2,

$$\frac{\partial x_3}{\partial x_{2,0}} > 0, t > 0.$$

Example 23 (Epidemiology): We consider again the SIS epidemic model of Example 9. The Jacobian matrix for this problem is

$$D\tilde{f} = \begin{pmatrix} -\beta x_2 & \gamma - \beta x_1 \\ \beta x_2 & -\gamma + \beta x_1 \end{pmatrix}.$$

The work of Chapter 2 gave us some results, but we must still prove that

$$\frac{\partial x_1}{\partial x_2(0)} < 0, t > 0, \frac{\partial x_2}{\partial x_1(0)} > 0, t > 0, \text{ and } \frac{\partial x_2}{\partial x_2(0)} > 0, t \geq 0.$$

Consider first the proper cone K_1 with extreme rays

$$\tilde{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \tilde{e}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

pictured in Figure 29. We calculate the images of the extreme rays under $D\tilde{f}$:

$$D\tilde{f}\tilde{e}_1 = \beta x_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \beta x_2 \tilde{e}_2, \text{ and} \tag{3.44}$$

$$D\tilde{f}\tilde{e}_2 = (\beta x_2 + \gamma - \beta x_1) \begin{pmatrix} -1 \\ 1 \end{pmatrix} = (\beta x_2 + \gamma - \beta x_1)\tilde{e}_2. \tag{3.45}$$

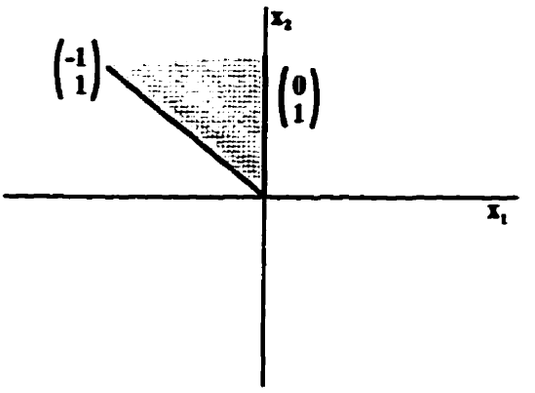


Figure 30: The proper (polyhedral) cone K_2 for Example 23.

K_1 is preserved by $D\tilde{f}$ and induces the graph

which, upon applying Corollary 37, gives

$$\frac{\partial x_2}{\partial x_1(0)} > 0, t > 0, \text{ and } \frac{\partial x_2}{\partial x_2(0)} > 0, t \geq 0.$$

Consider second the proper cone K_2 with extreme rays

$$\tilde{e}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ and } \tilde{e}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

drawn in Figure 30. Compared to K_1, one extreme ray has changed. The new extreme ray has image $(-\gamma + \beta x_1)\tilde{e}_2$, meaning that the same graph is induced by $D\tilde{f}$. Corollary 37 tells us that

$$\frac{\partial x_1}{\partial x_2(0)} < 0, t > 0.$$

Example 24 (Chemical Kinetics): We return to the Michaelis-Menten enzyme kinetics model of Example 10. For ease of reference, we present the table of partial derivative signs in Table 9. In the work of Example 10, we were only able to obtain

	$x_1(t)$	$x_2(t)$	$x_3(t)$	$x_4(t)$
$x_1(0)$	+	−	*	+
$x_2(0)$	−	+	+	+
$x_3(0)$	+	*	*	+
$x_4(0)$	0	0	0	+

Table 9: Behaviour of concentrations with respect to changes in initial concentrations for the Michaelis-Menten system.

the signs in the bottom row of Table 9; now, we will apply the work of this chapter. The system of ordinary differential equations for this problem is

$$\dot{x}_1(t) = f_1(x_1, x_2, x_3, x_4) = -k_1 x_1 x_2 + (k_{-1} + k_2) x_3, \tag{3.46}$$

$$\dot{x}_2(t) = f_2(x_1, x_2, x_3, x_4) = -k_1 x_1 x_2 + k_{-1} x_3, \tag{3.47}$$

$$\dot{x}_3(t) = f_3(x_1, x_2, x_3, x_4) = k_1 x_1 x_2 - (k_{-1} + k_2) x_3, \text{ and} \tag{3.48}$$

$$\dot{x}_4(t) = f_4(x_1, x_2, x_3, x_4) = k_2 x_3. \tag{3.49}$$

The Jacobian matrix for this problem is

$$D\tilde{f} = \begin{pmatrix} -k_1 x_2 & -k_1 x_1 & k_{-1} + k_2 & 0 \\ -k_1 x_2 & -k_1 x_1 & k_{-1} & 0 \\ k_1 x_2 & k_1 x_1 & -(k_{-1} + k_2) & 0 \\ 0 & 0 & k_2 & 0 \end{pmatrix}.$$

Notice that x_1, x_2, and x_3 do not depend on x_4; we can consider their three-dimensional subsystem. In this case, the 3×3 Jacobian matrix $D\tilde{f}$ will be the upper left 3×3 block of the Jacobian matrix for the full system.

Letting $M_1 = x_1(0) + x_3(0)$ and $M_2 = x_2(0) + x_3(0) + x_4(0)$, it is easily observable that $x_1(t) \leq M_1$ and $x_i(t) \leq M_2$, $i = 2, 3, 4$, $\forall\, t \geq 0$, since $\dot{x}_2 + \dot{x}_3 + \dot{x}_4 = 0$ and $\dot{x}_1 + \dot{x}_4 = 0$.

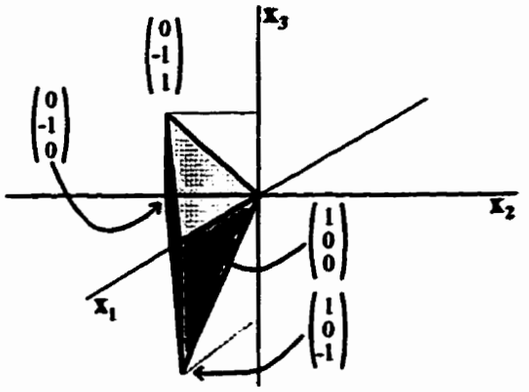


Figure 31: A proper (polyhedral) cone K_1 for Example 24.

Consider first the three-dimensional proper (polyhedral) cone K_1 in $\mathbb{R}^3$ with extreme rays

$$\tilde{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tilde{e}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \tilde{e}_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \text{ and } \tilde{e}_4 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix},$$

drawn in Figure 31. Notice that $\tilde{e}_1 + \tilde{e}_3 = \tilde{e}_2 + \tilde{e}_4$; by Proposition 41 (and as Figure 31 indicates), K_1 is a polyhedral cone. The four faces of K_1 are given by $F_{1,2} = sp^+\{\tilde{e}_1, \tilde{e}_2\}$, $F_{1,4} = sp^+\{\tilde{e}_1, \tilde{e}_4\}$, $F_{2,3} = sp^+\{\tilde{e}_2, \tilde{e}_3\}$, and $F_{3,4} = sp^+\{\tilde{e}_3, \tilde{e}_4\}$, where sp^+ denotes all non-negative combinations of the vectors listed. We calculate the images of these extreme rays under $D\tilde{f} + lI$, for appropriate choices of l.

$$(D\tilde{f} + k_1x_2)\tilde{e}_1 = k_1x_2 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = k_1x_2\tilde{e}_4,$$

$$(D\tilde{f} + (k_1x_1 + k_{-1} + k_2))\tilde{e}_2 = \begin{pmatrix} 0 \\ -k_1y - k_{-1} \\ 0 \end{pmatrix} = (k_1x_2 + k_{-1})\tilde{e}_3,$$

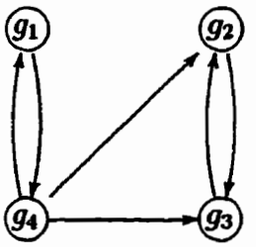


Figure 32: $G_{K_1}(\tilde{f}, \mathcal{O})$ for Example 24

$$(D\tilde{f} + k_1 x_1)\tilde{e}_3 = k_1 x_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = k_1 x_1 \tilde{e}_2, \text{ and}$$

$$(D\tilde{f} + (k_1 x_2 + k_{-1} + k_2))\tilde{e}_4 = \begin{pmatrix} k_1 x_2 + k_{-1} + k_2 \\ -k_2 \\ 0 \end{pmatrix}$$
$$= (k_1 x_2 + k_{-1} + k_2)\tilde{e}_1 + k_2 \tilde{e}_3.$$

Picking $l = k_1(M_1 + x_2(0)) + k_{-1} + k_2$, say, guarantees that K_1 is preserved by $D\tilde{f}$. Using the known face structure of K_1, we conclude that $D\tilde{f} + lI$ maps $\tilde{e}_1$ to $F_{1,4}$, $\tilde{e}_2$ to $F_{2,3}$, $\tilde{e}_3$ to $F_{2,3}$, and $\tilde{e}_4$ to the interior of K_1. The associated multigraph $G_{K_1}(\tilde{f}, \mathcal{O})$ is given in Figure 32. Applying Corollary 37, we can conclude that for $t > 0$

$$\frac{\partial x_1}{\partial x_1(0)} > 0, \text{ and } \frac{\partial x_2}{\partial x_1(0)} > 0.$$

Next consider the three-dimensional proper (polyhedral) cone K_2 in $\mathbf{R}^3$ with extreme rays

$$\tilde{e}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \tilde{e}_2 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \tilde{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \text{ and } \tilde{e}_4 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

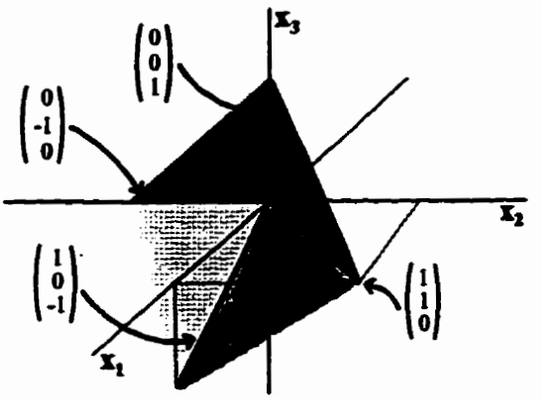


Figure 33: A proper (polyhedral) cone K_2 for Example 24.

presented in Figure 33. Notice that $\tilde{e}_1 + \tilde{e}_3 = \tilde{e}_2 + \tilde{e}_4$; by Proposition 41 (and as Figure 33 shows), K_2 is a polyhedral cone. The four faces of K_2 are given by $F_{1,2} = sp^+\{\tilde{e}_1, \tilde{e}_2\}$, $F_{1,4} = sp^+\{\tilde{e}_1, \tilde{e}_4\}$, $F_{2,3} = sp^+\{\tilde{e}_2, \tilde{e}_3\}$, and $F_{3,4} = sp^+\{\tilde{e}_3, \tilde{e}_4\}$, where sp^+ denotes all non-negative combinations of the vectors listed. Again, calculate the images of the extreme rays under $D\tilde{f} + lI$.

$$(D\tilde{f} + (k_1x_2 + k_{-1} + k_2))\tilde{e}_1 = \begin{pmatrix} 0 \\ -k_1y - k_{-1} \\ 0 \end{pmatrix} = (k_1x_2 + k_{-1})\tilde{e}_2,$$

$$(D\tilde{f} + k_1x_1)\tilde{e}_2 = k_1x_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = k_1x_1\tilde{e}_1,$$

$$(D\tilde{f} + k_{-1} + k_2)\tilde{e}_3 = \begin{pmatrix} k_{-1} + k_2 \\ k_{-1} \\ 0 \end{pmatrix} = (k_{-1} + k_2)\tilde{e}_4 + k_2\tilde{e}_2, \text{ and}$$

$$(D\tilde{f} + k_1(x_1 + x_2))\tilde{e}_4 = k_1(x_1 + x_2) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = k_1(x_1 + x_2)\tilde{e}_3.$$

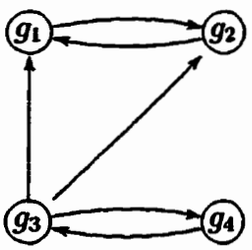


Figure 34: $G_{K_2}(\tilde{f}, \mathcal{O})$ for Example 24

Picking $l = k_1(M_1 + M_2) + k_{-1} + k_2$ guarantees that $D\tilde{f}$ preserves K_2 and, using the face structure of K_2, that the image under $D\tilde{f} + lI$ of $\tilde{e}_1$ is in $F_{1,2}$, the image of $\tilde{e}_2$ is in $F_{1,2}$, the image of $\tilde{e}_3$ is in the interior of K_2, and the image of $\tilde{e}_4$ is in $F_{3,4}$. $G_{K_2}(\tilde{f}, \mathcal{O})$ is presented in Figure 34. By Corollary 37, K_2 lets us conclude that for $t > 0$

$$\frac{\partial x_1}{\partial x_3(0)} > 0.$$

Notice from the work with K_1 that the two-dimensional cone with extreme rays

$$\tilde{e}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \text{ and } \tilde{e}_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

is preserved by $D\tilde{f}$. To avoid the minus signs, consider the two-dimensional proper (polyhedral) cone K_3 in $\mathbf{R}^3$ with extreme rays

$$\tilde{e}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \text{ and } \tilde{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

drawn in Figure 35. Repeating earlier work, the images under $D\tilde{f} + lI$ of the

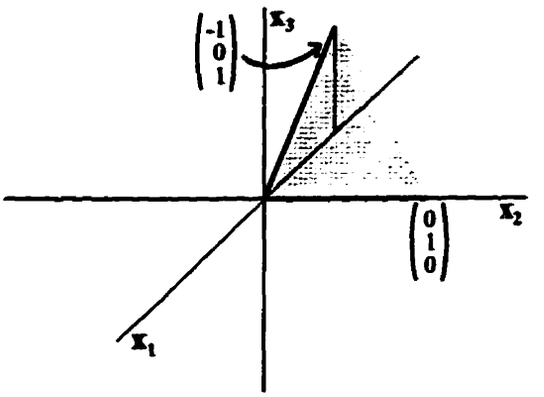


Figure 35: A proper (polyhedral) cone K_3 for Example 24.

extreme rays are

$$(D\tilde{f} + (k_1x_2 + k_{-1} + k_2))\tilde{e}_1 = \begin{pmatrix} 0 \\ k_1y + k_{-1} \\ 0 \end{pmatrix} = (k_1x_2 + k_{-1})\tilde{e}_2, \text{ and}$$

$$(D\tilde{f} + k_1x_1)\tilde{e}_2 = k_1x_1 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = k_1x_1\tilde{e}_1.$$

Once again, picking $l = k_1(M_1 + M_2) + k_{-1} + k_2$ guarantees that the extreme rays are mapped to the relative interior of K_3 by $D\tilde{f} + lI$, implying that the multigraph $G_{K_3}(\tilde{f}, \mathcal{O})$ is strongly connected. Applying Corollary 37 allows us to conclude that for $t > 0$

$$\frac{\partial x_1}{\partial x_2(0)} < 0, \frac{\partial x_2}{\partial x_2(0)} > 0, \text{ and } \frac{\partial x_3}{\partial x_2(0)} > 0.$$

The partial derivative results for x_4 will require cones in four dimensions, adding new difficulties because we can no longer picture things. There is future work to be done investigating the face structure of cones in higher dimensions; we will see where this comes in. Luckily, for this problem a simplification occurs.

Consider the proper (polyhedral) cone K_4 with extreme rays

$$\tilde{e}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \tilde{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \text{ and } \tilde{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

K_4 is a three-dimensional cone in $\mathbb{R}^4$. The images of the extreme rays under $D\tilde{f}+lI$ are given by

$$(D\tilde{f} + k_1x_2 + k_{-1} + k_2)\tilde{e}_1 = \begin{pmatrix} 0 \\ k_1x_2 + k_{-1} \\ 0 \\ k_2 \end{pmatrix} = (k_1x_2 + k_{-1})\tilde{e}_2 + k_2\tilde{e}_3,$$

$$(D\tilde{f} + k_1x_1)\tilde{e}_2 = k_1x_1 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = k_1x_1\tilde{e}_1, \text{ and}$$

$$(D\tilde{f})\tilde{e}_3 = \tilde{0}.$$

Pick $l = k_1(M_1+M_2)+k_{-1}+k_2$ to guarantee that $D\tilde{f}$ preserves K_4. As it turns out, in this case we can draw the multigraph $G_{K_4}(\tilde{f}, \mathcal{O})$ because the face structure of K_4 is reasonably simple to see: the three extreme rays are the one-dimensional faces; there are three two-dimensional faces, each consisting of non-negative combinations of a pair of extreme rays; and there is one three-dimensional face consisting of the non-negative combinations of the three extreme rays. $G_{K_4}(\tilde{f}, \mathcal{O})$ given in Figure 36. Applying Corollary 37, we can conclude that

$$\frac{\partial x_4}{\partial x_2(0)} > 0, \text{ for } t > 0.$$

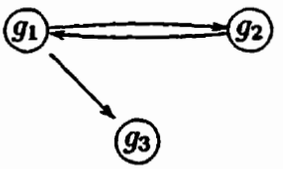


Figure 36: $G_{K_4}(\tilde{f}, \mathcal{O})$ for Example 24

We consider the proper (polyhedral) cone K_5 with extreme rays

$$\tilde{e}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \tilde{e}_2 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \tilde{e}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \text{ and } \tilde{e}_4 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

Notice that $\tilde{e}_1 + \tilde{e}_4 = \tilde{e}_2 + \tilde{e}_3$; by Proposition 41, K_5 is a polyhedral cone. The four faces of K_5 are given by $F_{1,2} = sp^+\{\tilde{e}_1, \tilde{e}_2\}$, $F_{1,3} = sp^+\{\tilde{e}_1, \tilde{e}_3\}$, $F_{2,4} = sp^+\{\tilde{e}_2, \tilde{e}_4\}$, and $F_{3,4} = sp^+\{\tilde{e}_3, \tilde{e}_4\}$, where sp^+ denotes all non-negative combinations of the vectors listed. The images of the extreme rays under $D\tilde{f} + lI$ are given by

$$(D\tilde{f} + k_{-1} + k_2)\tilde{e}_1 = \begin{pmatrix} k_{-1} + k_2 \\ k_{-1} \\ 0 \\ k_2 \end{pmatrix} = (k_{-1} + k_2)\tilde{e}_3 + k_2\tilde{e}_2, \quad (3.50)$$

$$(D\tilde{f} + k_1x_1)\tilde{e}_2 = k_1x_1 \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} = k_1x_1\tilde{e}_4, \quad (3.51)$$

$$(D\tilde{f} + k_1(x_1 + x_2))\tilde{e}_3 = k_1(x_1 + x_2) \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

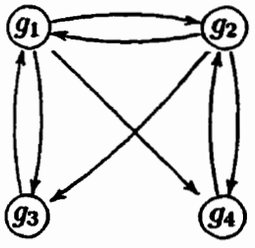


Figure 37: $G_{K_5}(\tilde{f}, \mathcal{O})$ for Example 24

$$= k_1(x_1 + x_2)\tilde{e}_1, \text{ and} \tag{3.52}$$

$$(D\tilde{f} + k_1x_2 + k_{-1} + k_2)\tilde{e}_4 = (k_1x_2 + x_{-1}) \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

$$= (k_1x_2 + x_{-1})\tilde{e}_2. \tag{3.53}$$

Certainly, then, for $l = k_1(M_1 + M_2) + k_{-1} + k_2$, K_5 is preserved by $D\tilde{f}$. Since we are lucky enough to know the face structure of K_5, we can see that $D\tilde{f} + lI$ maps $\tilde{e}_1$ and $\tilde{e}_2$ into the relative interior of K_5. $D\tilde{f} + lI$ maps $\tilde{e}_3$ to $F_{1,3}$ and maps $\tilde{e}_4$ to $F_{2,4}$. $G_{K_5}(\tilde{f}, \mathcal{O})$ is drawn in Figure 37. Since K_5 contains the x_3−axis, $x_4 \geq 0$ in K_5, and g_2 is strongly connected to all other vertices in $G_{K_5}(\tilde{f}, \mathcal{O})$, by Corollary 37, we can conclude that

$$\frac{\partial x_4}{\partial x_3(0)} > 0, \text{ for } t > 0.$$

The final result we wish to obtain is

$$\frac{\partial x_4}{\partial x_1(0)} > 0, \ t > 0.$$

We consider the proper (polyhedral) cone K_6 with extreme rays

$$\tilde{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tilde{e}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \tilde{e}_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \text{ and } \tilde{e}_4 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

In this case, $\tilde{e}_1 + \tilde{e}_3 = \tilde{e}_2 + \tilde{e}_4$; by Proposition 41, K_6 is a polyhedral cone. The four faces of K_6 are given by $F_{1,2} = sp^+\{\tilde{e}_1, \tilde{e}_2\}$, $F_{1,4} = sp^+\{\tilde{e}_1, \tilde{e}_4\}$, $F_{2,3} = sp^+\{\tilde{e}_2, \tilde{e}_3\}$, and $F_{3,4} = sp^+\{\tilde{e}_3, \tilde{e}_4\}$. The images of the extreme rays under $D\tilde{f} + lI$ are given by

$$(D\tilde{f} + k_1x_2)\tilde{e}_1 = k_1x_2 \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} = k_1x_2\tilde{e}_2, \tag{3.54}$$

$$(D\tilde{f} + k_1x_1 + k_{-1} \div k_2)\tilde{e}_2 = \begin{pmatrix} k_1x_1 + k_{-1} + k_2 \\ -k_2 \\ 0 \\ k_2 \end{pmatrix} \tag{3.55}$$

$$= (k_1x_1 + k_{-1} + k_2)\tilde{e}_1 + k_2\tilde{e}_3, \tag{3.56}$$

$$(D\tilde{f} + k_1x_1)\tilde{e}_3 = k_1x_1 \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} = k_1x_1\tilde{e}_4, \text{ and} \tag{3.57}$$

$$(D\tilde{f} + k_1x_2 + k_{-1} + k_2)\tilde{e}_4 = (k_1x_2 + x_{-1}) \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} = (k_1x_2 + x_{-1})\tilde{e}_3. \tag{3.58}$$

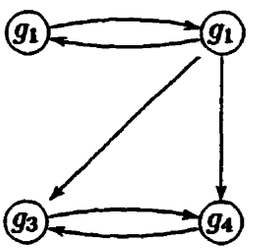


Figure 38: $G_{K_6}(\tilde{f}, \mathcal{O})$ for Example 24

Again, for $l = k_1(M_1 + M_2) + k_{-1} + k_2)$, K_6 is preserved by $D\tilde{f}$. This time around, we see that $D\tilde{f} + lI$ maps $\tilde{e}_1$ to $F_{1,2}$, $\tilde{e}_2$ to $relint(K_6)$, $\tilde{e}_3$ to $F_{3,4}$, and $\tilde{e}_4$ to $F_{3,4}$, We draw $G_{K_6}(\tilde{f}, \mathcal{O})$ in Figure 38. Since K_6 contains the x_1-axis, $x_4 \geq 0$ in K_6, and g_1 is strongly connected to all other vertices in $G_{K_6}(\tilde{f}, \mathcal{O})$, by Corollary 37, we can conclude that

$$\frac{\partial x_4}{\partial x_1(0)} > 0, \text{ for } t > 0.$$

Chapter 4

Directions for Future Work

In [18], chain reactions of the form

$$A+B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C_1 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} C_2 \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} \cdots \underset{k_{-(n-1)}}{\overset{k_{n-1}}{\rightleftharpoons}} C_{n-1} \underset{k_{-n}}{\overset{k_n}{\rightleftharpoons}} C_n$$

and

$$A+B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C_1 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} C_2 \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} \cdots \underset{k_{-n}}{\overset{k_n}{\rightleftharpoons}} C_n \overset{k_{n+1}}{\rightharpoonup} D$$

were analyzed. In each case, the arguments were quite lengthy and involved. It would be interesting to attempt to apply the methods of Chapter 3 to these general problems.

Of course, the face structure of any cones that might seem helpful for these problems would need to be known: a second topic for future work is monotonicity with respect to cones in higher dimensions. Understanding face structure is crucial to the final theorems in this thesis.

At this stage, we have no sufficient conditions for strong monotonicity with respect to non-polyheral proper cones.

It would also be nice to find a necessary condition for strong monotonicity. As we saw in Example 17, the sufficient condition that has proved useful in this work is not necessary.

A careful proof of Proposition 35 should be developed.

There are similar graphs to those in Section 3.2.2 ($G_{K,p}(\tilde{f}, N)$) in the literature ([2],[3]). An exposition of these graphs would prove interesting.

Bibliography

[1] G. Aronsson and R.B. Kellogg. On a differential equation arising from compartmental analysis. *Math. Biosci.*, 38:113–122, 1978.

[2] G. Barker and B. Tam. Graphs for cone preserving maps. *Linear Algebra and its Applications*, 37:199–204, 1981.

[3] A. Berman. Convexity, graph theory and non-negative matrices. *Annals of Discrete Mathematics*, 20:55–59, 1984.

[4] A. Berman, M. Neumann, and R. Stern. *Nonnegative Matrices in Dynamic Systems.* John Wiley and Sons, New York, 1989.

[5] A. Berman and R.J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences.* Academic Press, New York, 1979.

[6] H. Edelsbrunner. *Algorithms in Combinational Geometry.* Springer-Verlag, Berlin, 1987.

[7] L. Edelstein-Keshet. *Mathematical Models in Biology.* Random House, New York, 1988.

[8] P. Erdi and J. Toth. *Mathematical Models of Chemical Reactions.* Princeton University Press, New Jersey, 1989.

[9] F.R. Gantmacher. *The Theory of Matrices*, volume 2. Chelsea, New York, 1964.

[10] G. Hadley. *Linear Programming.* Addison-Wesley, Reading, 1962.

[11] M.W. Hirsch. Differential equations and convergence almost everywhere in strongly monotone semiflows. *Contemporary Mathematics*, 17:267–285, 1983.

[12] M.W. Hirsch. The dynamical systems approach to differential equations. *Bull. Amer. Math. Soc.*, 11(1):1–64, 1984.

[13] M.W. Hirsch. Systems of differential equations that are competitive or cooperative II: Convergence almost everywhere. *SIAM J. Math. Anal.*, 16(3):423–439, 1985.

[14] M.W. Hirsch. Network dynamics: Principles and problems. In F. Pasemann and H.D. Doebner, editors, *Neurodynamics*, pages 3–29, Singapore, 1991. World Scientific.

[15] C. Jeffries. Qualitative stability and digraphs in model ecosystems. *Ecology*, 55:1415–1419, 1974.

[16] C. Jeffries. *Code Recognition and Set Selection with Neural Networks.* Birkhäuser, Boston, 1991.

[17] C. Jeffries, V. Klee, and P. Van Den Driessche. When is a matrix sign stable? *Can. J. Math.*, 29(2):315–326, 1977.

[18] H. Kunze. Monotonicity properties of systems of ordinary differential equations. Master's thesis, University of Waterloo, April 1992.

[19] H. Kunze and D. Siegel. A graph theoretical approach to monotonicity with respect to initial conditions. In X. Liu and D. Siegel, editors, *Comparison Methods and Stability Theory*, pages 207–216, New York, 1994. Marcel Dekker.

[20] H. Kunze and D. Siegel. Chemical reactions which induce an order preserving flow, to appear.

[21] H. Kunze and D. Siegel. A graph theoretical approach to monotonicity with respect to initial conditions II, to appear.

[22] G.S. Ladde, V. Lakshmikantham, and A.S. Vatsala. *Iterative Techniques for Nonlinear Differential Equations.* Pitman, Boston, 1985.

[23] V. Lakshmikantham and S. Leela. Cone-valued lyapunov functions. *Nonlinear Analysis*, 1:215–222, 1977.

[24] R. Martin. Asymptotic stability and critical points for nonlinear quasimonotone parabolic systems. *J. Diff. Eqns.*, 30:391–423, 1978.

[25] F.P. Perparata and M.I. Shamos. *Computational Geomtery.* Springer-Verlag, New York, 1985.

[26] J.P. Quirk and R. Ruppert. Qualitative economics and the stability of equilibrium. *Rev. Econ. Stud.*, 32:311–326, 1965.

[27] A. Recski. *Matroid Theory and its Applications.* Springer-Verlag, Hungary, 1989.

[28] R.M. Redheffer and W. Walter. A differential inequality for the distance function in normed linear spaces. *Math. Ann.*, 211:299–314, 1974.

[29] R.M. Redheffer and W. Walter. Flow-invariant sets and differential inequalities in normed spaces. *Applicable Analysis*, 5:149–161, 1975.

[30] D. Siegel. Monotonicity properties of solutions to systems of differential equations arising in chemical kinetics. In C.M. Dafermos, G. Ladas, and G. Papanicolaou, editors, *Differential Equations*, pages 637–646, New York, 1989. Marcel Dekker.

[31] D. Siegel and H. Kunze. Monotonicity properties of solutions to the SIS and SIR epidemic models. *J. Math. Anal. Appl.*, 185(1):65–85, July 1994.

[32] D. Siegel and D.W. Lozinski. Monotonicity properties of the Michaelis-Menten reactions of enzyme kinetics. *Rocky Mountain J. Math.*, 20(4):1157–1172, 1990.

[33] W. Slomzxynski. Irreducible cooperative systems are strongly monotone. *Univ. Iagel. Acta. Math.*, 30:87–113, 1993.

[34] H.L. Smith. Systems of ordinary differential equations which generate an order preserving flow. *SIAM Rev.*, 30(1):87–113, March 1988.

[35] H.L. Smith. *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems.* American Mathematical Society, Providence, 1995.

[36] J. Vandergraft. Spectral properties of matrices which have invariant cones. *SIAM J. Appl. Math.*, 16:1208–1222, 1968.

[37] P. Volkmann. Gewöhnliche Differentialgleichungen mit quasimonoton wachsenden Funktionen in topologischen Vektorräumen. *Math. Z.*, 127:157–164, 1972.

[38] P. Volkmann. Über die Invarianz konvexer Mengen und Differentialgleichungen in einem normierten Raum. *Math. Ann.*, 203:201–210, 1973.

[39] A.I. Vol'pert and S.I. Khudyaev. *Analysis in Classes of Discontinuous Functions and Equations of Mathematical Physics.* Marinus Nijhoff, Dordrecht, Netherlands, 1985.

[40] W. Walter. Ordinary differential inequalities in ordered banach spaces. *J. Differential Equations*, 9:253–261, 1971.

[41] W. Walter. On strongly monotone flows. *Annales Polonici Mathematici*, to appear.

[42] R. Webster. *Convexity.* Oxford Science Publications, Oxford, 1994.

[43] Y. Wong and K. Ng. *Partially Ordered Topological Vector Spaces.* Clarendon Press, Oxford, 1973.

[44] N.V. Yefimov. *Quadratic Forms and Matrices.* Academic Press, New York, 1964.

[45] G.M. Ziegler. *Lectures on Polytopes.* Springer-Verlag, New York, 1995.